Hereditary Haemorrhagic Telangiectasia:

Linkage mapping and identification of the second locus.

Jonathan N. Berg

M.D.

University of Edinburgh

1998

This thesis is dedicated to the people with Hereditary Haemorrhagic Telangiectasia.

Acknowledgements

I wish to thank the following people

My Supervisor, Dr. Mary Porteous for the teaching, support and encouragement, without which this work would never have taken place.

Dr. Doug Marchuk for inviting and funding me to work in his team at Duke University, and his supervision and hospitality during my stay.

Carol Gallione (and Paul) for technical assistance, endless patience, and introducing me to real American life.

Andreas and Gabi for ideas, friendship and the best German cooking in North Carolina.

Gene Jackson for help with family 9.

Jon Warner, Austin Diamond, Dave Johnson, Tim Stenzel and Alan Guttmacher along with many others for teaching, discussion and assistance.

The Scottish Hospitals Endowment Research Fund for the Jean V. Baxter Fellowship allowing me to complete this work and move onto the next.

And finally Alison, who waited for me to come back.

Declaration.

This thesis has been composed solely by myself.

The data in this thesis form part of a collaboration with the research team of Dr. Doug Marchuk, Associate Professor in the Department of Genetics at Duke University, North Carolina. I wish to acknowledge the following contributions:

Dave Johnson and Carol Gallione ran the microsattelite markers used for the American linkage analysis data and haplotyping.

Tim Stenzel identified the borders of exons 7,8,9 and 10 of ALK-1, and performed the mutation analysis for exons 9 and 10.

The work in this thesis is otherwise my own.

Jonathan Berg

Abstract

Hereditary Haemorrhagic telangiectasia (HHT) is an autosomal dominant disease which is characterised by frequent, severe nose bleeds, muco-cutaneous telangiectases, gastro-intestinal haemorrhage and the formation of visceral arteriovenous malformations, particularly in the pulmonary and cerebral circulation. The endoglin gene on chromosome 9q34 was shown to be mutated in a proportion of patients with this disease. Endoglin is a Transforming Growth Factor-beta (TGF-β) binding protein expressed predominantly by endothelial cells. The endoglin locus (HHT1) had previously been excluded in some families by linkage, suggesting that mutations in a second unidentified gene caused HHT in a proportion of patients.

The work presented in this thesis aimed to identify the second gene for HHT (HHT2) and compare features of patients with mutations in this gene, and those with mutations in Endoglin.

Members of 5 new families who suffer from HHT were assessed. Three of these families show exclusion of the chromosome 9 locus on linkage analysis. Linkage analysis of these 3 families in conjunction with analysis of other families previously described demonstrated a second locus for HHT on chromosome 12. A candidate interval of approximately 1cM was defined by haplotype analysis, looking at critical recombination events. A strong candidate gene was identified within this region. This gene, ALK-1, is a type I serine-threonine kinase receptor with expression demonstrated solely in endothelial cells.

A large genomic clone containing the ALK-1 was identified from a PAC library. The genomic structure of ALK-1 was elucidated. Conditions for the PCR amplification of each exon of the ALK-1 from genomic DNA gene were determined, allowing sequencing and mutation analysis of the gene in patients with HHT. The ALK-1 gene was sequenced in 12 patients. Mutations in ALK-1 were identified in all 6 patients from families with evidence of linkage to 12q or exclusion of linkage to 9q34. Three mutations were also identified in 6 patients from whom no linkage data was available. The mutations identified left the transmembrane domain of the ALK-1 receptor intact, but are predicted to disrupt either kinase function, or ligand binding to the extracellular domain. Whether expression of an abnormal protein is necessary for disease causation, or whether mutations lead to a null allele remains to be established. The identification of ALK-1 as a second gene responsible for HHT is the first evidence that there is an interaction between ALK-1 and Endoglin in the endothelial cell.

The comparison of families with known endoglin mutations and ALK-1 mutations showed a much higher incidence of pulmonary arterio-venous malformations in patients with endoglin mutations. The incidence of gastro-intestinal haemorrhage and hepatic complications was not significantly different. The reason for this remains to be established.

Contents

Chapter 1. Introduction

A. Hereditary Haemorrhagic Telangiectasia, - a historical perspective.

The original name of the disorder, Osler-Rendu-Weber syndrome originates from three physicians who gave early descriptions of the disease. Osler (1849 to 1919), a prominent American physician, described the disorder in 1901 in his paper "Family form of recurring epistaxis associated with multiple telangiectases of the skin and mucous membranes" . In this paper he referred to an earlier description of the disease, written by the French physician Rendu in 1896.

The paper by Rendu, "Epistaxis repetees chez un sujet porteur de petits angiomes cutanes et muqueux." (recurring nosebleeds in a patient with small cutaneous angiomas of the skin and mucous membranes.) remains one of the best descriptions of the disease. Rendu described a 52 year old man admitted with a prolonged nosebleed. He has a history of epistaxis occurring several times a day, which started at the age of 12 becoming much more frequent after the age of 35. He was not, however prone to abnormal bruising, bleeding from the gums, or prolonged bleeding after an injury. He was anaemic and mildly jaundiced, although showing no overt signs of liver failure. The telangiectases are described:

"Il existe sur la peau du nez, des joues, de la levre superieure et du menton de petites taches pourprees ...les plus grosses atteignant les dimensions d'une lentille, et qui sont de veritables angiome cutanes, produit par une dilatation des vaisseaux superficiels de la peau..."

On the skin of the nose, cheeks, upper lip and chin, there are small purple marks.....the largest approach the size of a lentil, and these are true cutaneous angiomas produced by a dilatation of the superficial vessels of the skin....

The distribution of these telangiectases on the face, lips and mucous membranes, sparsely on the trunk but not on the limbs is noted, as is their tendency to blanche incompletely and refill quickly after ceasing pressure on them. The patient's father died of gastro-intestinal haemorrhage, and his brother, who died of other causes, suffered from frequent nosebleeds.

Parkes-Weber was an English physician had a lifelong interest in vascular lesions wrote an article in the Lancet in 1907 describing a mother and 9 children, noting that the features in some of the children, even after the age of 30 were so subtle as to make diagnosis difficult. He also reviewed previous descriptions by Legg in 1876 and Chiari in 1887 of diseases similar to Osler-Rendu-Weber syndrome and of a family with a history of nosebleeds reported by B.G. Babbington in the Lancet in 1865.

Earlier descriptions may well have described Osler-Rendu-Weber syndrome. H.G. Sutton in the Medical Mirror in 1864 described a possible association between haemoptysis in later life (presumed to be due to *the phthisis* or TB) and epistaxis. He cites a case described by Sir Henry Marsh of a 46 year old lady with haemoptysis. Her father who had frequent epistaxis died following a haemoptysis. Her brother and 3 of her sons had frequent nosebleeds, whereas her daughter did not. This association between epistaxis and haemoptysis was well recognised long before this. Referring to this association in a dissertation presented to the Royal Medical Society in Edinburgh in 1800, C.Rattray wrote:

"Haemoptysis, or the spitting of blood from the lungs...This disease is frequently observed in those of the most sensible and irritable disposition, of a weak slender make, with long necks and flat breasts and those in any way malformed about the thorax, those in their youth have been subject to epistaxis..."

The more descriptive name "Hereditary Haemorrhagic Telangiectasia" originated in a comprehensive discussion of the disorder written by Hanes in 1909, while a resident at Johns Hopkins Hospital. This name is descriptive and avoids giving insult to the shades of many early physicians who described the disease, but who are not remembered in the

eponymous name. Throughout this text, therefore, the name Hereditary Haemorrhagic Telangiectasia (abbreviated to HHT) will be used.

B. Clinical Description of Hereditary Haemorrhagic Telangiectasia.

HHT is an autosomal dominant disease characterised by recurrent epistaxis, muco-cutaneous telangiectasis and arteriovenous malformations, particularly in the lungs and brain. Vascular complications also affect the hepatic and gastro-intestinal systems, with involvement of the genito-urinary system and heart reported more rarely.

i.Incidence of HHT.

Two studies have attempted to arrive at an estimate of the incidence of HHT based on epidemiological recruitment of cases. A study of the French population (Bideau et.al.1980,1992) estimated an overall prevalence of 1 in 8,345 people. However, several areas had a much higher local prevalence, such as the Jura region (1 in 5062) and Deux-Sevres (1 in 4,287). This is in considerable contrast to estimates of between 1in 40,000 and 1 in 100,000 made prior to this by other authors but not based on an epidemiologically recruited study.

The point prevalence in the North of England was estimated at 1 in 39,216 based on questionnaires sent to general practitioners, ENT surgeons and

haematologists, revealing a minimum reported incidence of 79 cases in a population of 3.1 million (Porteous et.al 1992).

However, given the wide variation of severity of HHT, with its potential for under-diagnosis by the medical profession, it is likely that many cases in the population remain unidentified. The true prevalence may therefore be considerably higher than 1 in 40,000. A study in Vermont based on known cases of HHT, estimated the frequency in the Vermont population at about 1 in 16,500 (Guttmacher et.al. 1994).

Even though it is considered a rare disorder, with a prevalence of 1 in 40,000, over 1,000 people in the UK suffer from HHT. If the higher estimates are accurate, this figure may rise to over 5,000 affected individuals.

ii.Mucocutaneous Telangiectases.

The appearance of these is classically as described by Rendu (1896):

"...small purple marks, the size of a pinhead with the largest attaining the size of a lentil, which are true cutaneous angiomas, produced by a dilatation of the superficial vessels of the skin. They blanche on pressure, but do not completely disappear, blood returning immediately that pressure ceases."

As previously described elsewhere by several authors (Osler 1907, Bird et.al. 1957), the telangiectases can become raised and nodular, usually in later life. Spider angiomas have been reported (Bird et.al. 1957) but are not as characteristic, as these may have other causes.

Figure 1.1 shows typical mucocutaneous telangiectases.

A study in the French population found telangiectases in 74% of 240 patients on careful examination (Plauchu et.al.1989). About half of all patients who could remember the time of onset of lesions reported first appearance before the age of 30. They reported the most common site of telangiectases as the face (63% of patients) followed by the mouth (48%) and hands (37%). Lesions on the trunk, arms and lower limbs were much less common (less than 10% of patients) with many fewer lesions at these sites. For patients with oral involvement, the lower lip and tongue were most frequently involved. The palate was involved in 13% of patients as were the ears. The floor of mouth, conjunctiva, eyelids and gums were involved in 5% or fewer of patients.

In a British study, (Porteous et.al. 1992) the sites most frequently affected with telangiectases were reported as the palm and nailbed (71% of all patients), lips and tongue (66%) and the face (20%).

Figure 1.1 Typical muco-cutaneous lesions on the lips of a patient with hereditary haemorrhagic telangiectasia. This patient also has pulmonary arteriovenous malformations and cerebral arteriovenous malformations, in addition to daily heavy nosebleeds.

iii Epistaxis.

Epistaxis is usually the first symptom of HHT to be described, although the frequency and severity are very variable. Some people are only ever mildly affected, whereas in others epistaxis can be prolonged and at times, life threatening. Plauchu (Plauchu et.al.1989) found more than half his patients had epistaxis before the age of 20, with 90% complaining of the symptom by age 45. 70% of those between the ages of 45 and 60 had heavy episodes lasting more than 10 minutes.

Porteous (Porteous et.al. 1992) reported that out of 40 patients over the age of 50, 16 said the epistaxis had stayed the same, 16 said it had deteriorated and 8 said it had improved.

In a study of a population of patients screened for pulmonary arteriovenous malformations (Aasar et.al. 1991), an increase in frequency of epistaxis with age was seen, increasing from an average of 8 per month under age 20, to 22 per month over age 40.

Despite its frequency, epistaxis is one of the most difficult symptoms of HHT to treat. Local treatment such as cautery seems to have limited success (Porteous et.al.1992), although laser treatment may prove better. Hormonal treatment with oestrogen and progesterone has been advocated, but reports of its efficacy vary. It is also of limited value in

treating male patients. (McCaffrey et.al. 1977, Harrison et.al 1982, Vase et.al. 1981). Nasal septo-dermoplasty is one of several operative approaches that have been used, but after initial improvement of symptoms, epistaxis tends to recur. (McCaffrey et.al.1977).

iv.Gastro-Intestinal Haemorrhage.

This complication was less commonly encountered in the population based studies and usually starts in later life. 16% of patients seen in the French study (Plauchu et.al. 1989) were affected, but half the patients with this complication were over 58, only 1.5% of those under the age of 30 had gastro-intestinal haemorrhage. Porteous et.al reported 16 patients out of 56 suffering this complication, mainly from the upper gastro-intestinal tract.

A Danish study (Vase and Grove 1986) found gastro-intestinal haemorrhage in 35 out of 139 patients, only 2 of whom were under 40. Median age of onset was 55 compared to a median age of 11 for the onset of epistaxis. These patients had predominantly gastric and duodenal telangiectases, although 5 also had colonic telangiectases.
Various treatments for gastro-intestinal haemorrhage are available. Laser treatment can be effective if lesions are few in number (Bown et.al.1985), but the long term results may be less promising (Naveau et.al.1990). Desmopressin has been used with success to control acute bleeds refractory to normal management (Quitt et.al. 1990) Successful treatment

of those with severe haemorrhage using oestrogens is now established, (Van Cutsem et.al.1988,1990, Van Cutsem and Piessevaux 1996). The current regime uses 0.05mg ethinyloestradiol and 1mg norethisterone for 6 months. The reduction in haemorrhage is durable after treatment has been withdrawn after several months (Van Cutsem and Piessevaux 1996). Currently this seems the best first line treatment for those with refractory gastro-intestinal haemorrhage and transfusion dependence, when endoscopic treatment is not possible and the side effects of oestrogen treatment can be tolerated.

v. Hepatic Complications

The incidence of symptomatic hepatic lesions is low. In the french cohort, (Plauchu et.al.1989) 8% of patients had either hepatomegaly, jaundice, cirrhosis or a hepatic murmur. Onset of symptoms was usually in later life. Several families have been reported to have a much higher incidence of hepatic involvement than expected. A family with a father, 3 daughters and a grandson who all had symptomatic liver lesions with other classical features of HHT has been reported (Martini 1978). His review of all autopsy evidence to date suggested that fibrovascular lesions were a common finding, with evidence of cirrhosis in some patients. Angiographic findings in those with symptomatic hepatic lesions usually show diffuse small arteriovenous communications, precluding treatment by transcutaneous embolisation. A further small family of Northern Italian origin with

predominantly hepatic manifestations is reported by Ninkowictz (personal communication). A large Italian family has been described (Piantinada et.al.1996). Occasionally liver transplant has proved the only available treatment. (Bauer et.al. 1995).

vi Ocular Involvement

Conjunctival telangiectasia is not uncommon in HHT, affecting up to a third of patients (Brant et.al.1989). Blood stained tears have also been reported. (Wolper et.al.1969). Retinal telangiectases have not been well studied frequently, but haemorrhage from presumed retinal lesions have been reported. (Davis et.al.1971) The incidence of retinal lesions reported varies between by 10% (Brant et.al 1989) and 1/250 (Forker et.al 1963). This may represent a difference of ascertainment, or of the population studied.

vii Pulmonary Complications.

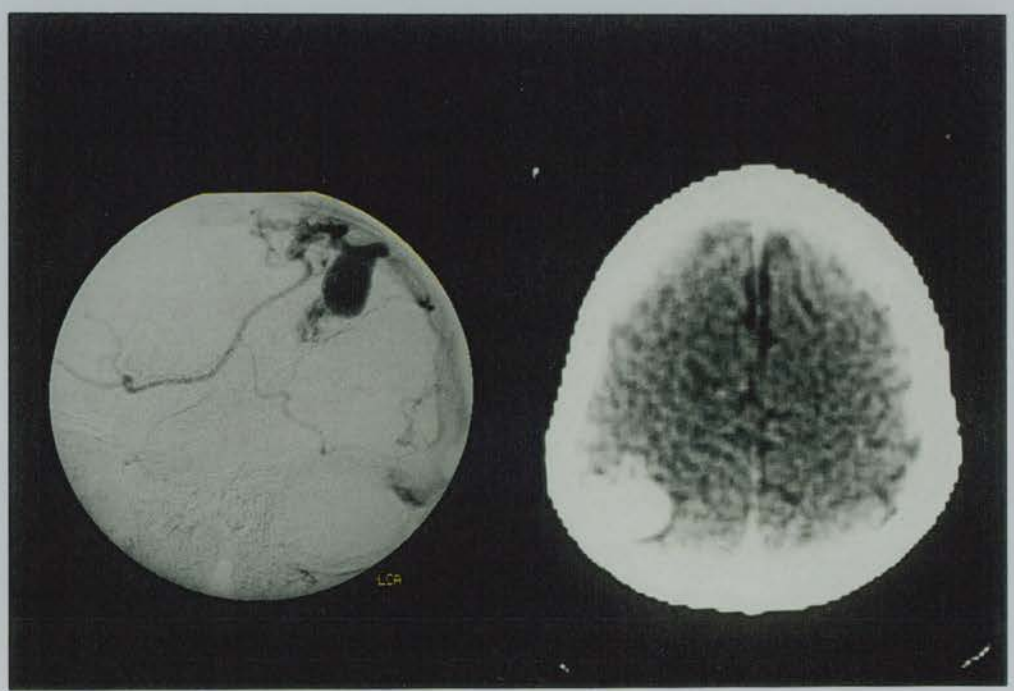
Pulmonary Arteriovenous Malformations (PAVMs) are one of the most characteristic lesions of HHT. They vary in size from insignificantly small to very large, and may be single or very numerous. An angiogram showing a PAVM is shown in **figure 1.2**.

Different reports give very varied estimates of incidence of this complication. Plauchu (Plauchu et.al. 1989) estimated that 4.6% of his population had PAVMs. In the North of England, (Porteous et.al.1992)

Figure 1.2 Characteristic arteriovenous malformations associated with hereditary haemorrhagic telangiectasia.

a. A pulmonary angiogram showing an unusually large pulmonary arteriovenous malformation in the right middle lobe. Multiple feeding vessels are visible.

b. A cerebral CT scan with contrast and a left carotid angiogram showing an arteriovenous malformation on the left hand side.



PAVMs were reported in 13/56 individuals. A Danish study (Vase et.al. 1985) found 16/95 patients had PAVMs on screening by chest X-ray. In a study of patients and their first degree relatives in Holland, (Haitjema et.al. 1995a,b) PAVMs were seen in 12/36 patients. Some early case reports found no PAVMs in families (Bird et.al. 1957, Hodgson et.al.1959) whereas a high incidence has been seen in other families (Hodgson et.al. 1959, Shovlin et.al.1994). This difference observed between families seemed too great to explain solely by differences in study technique. As discussed later, it was proposed that there are two forms of HHT, one with a higher incidence of visceral arteriovenous malformations.

Complications of PAVMs can arise as a result of a large pulmonary shunt. This manifests as cyanosis and hypoxia with finger clubbing and exercise limitation. More commonly, PAVMs are assymptomatic until they present with haemoptysis, stroke or cerebral abscess.

Haemoptysis or haemothorax has been reported as a complication in 11/143 patients with PAVMs and HHT (Ferrence et.al.1994). Bleeding in some of these cases may have come from endobronchial telangiectases (Wirth et.al. 1996). In an unpublished study (Charnas personal communication) a series of 102 patients referred for embolisation of PAVMs were screened by CT or MRI. 9 patients had clinically silent infarctions, 13 had subtle focal neurological findings and 21 had a history

of stroke confirmed on scanning. PAVMs may therefore have been causative of cerebral infarcts in up to 43 out of 102 patients. In 11 patients the stroke pre-dated the diagnosis of PAVM. Cerebral abscesses had affected 9 of this patient series.

Presence of a PAVM also confers a higher risk of complication in pregnancy, although it is difficult to accurately quantify this. Worsening of the pulmonary shunt is well documented and improves to some extent after delivery (Swinburn et.al. 1986). There is also a high incidence of stroke and pulmonary haemorrhage documented in women who become pregnant who are known to have a pre-existing PAVM. A series of 161 pregnancies in 47 women with HHT (Shovlin et.al. 1996) showed worsening shunt in 6 patients, stroke in 3 and fatal pulmonary haemorrhage in 2. In all but one case, complications occurred in the 12 mothers with known PAVMs, suggesting a much lower risk in mothers who have screened negative for a PAVM.

Screening of patients with HHT for assymptomatic PAVMs is advocated by most authors. However, not all authors agree on the best method of screening. Chest X-ray and pulse oximetry, measuring the saturation drop between supine and erect positioning has been advocated as a simple test (Hughes 1994), but never truly validated (Kjeldsen et.al.1996). The use of lung perfusion scan, when a right to left shunt would lead to increased

deposition of radiolabelled macro-aggregates in the kidneys may prove to be a more useful test (White et.al.1992). Measurement of arterial oxygen saturation or right to left shunt have been shown to be both sensitive and specific as screening tests. A shunt fraction of over 5% was seen in all except one of 35 patients with known PAVMs (Haitjema et.al.1995). Contrast echo-cardiography has been advocated as a sensitive screening test, although it remains to be fully assessed in this role.(Kjeldsen et.al 1996). Helical CT scanning is expensive but allows the demonstration of architecture of complex lesions. Angiography remains the gold standard for investigation of suspected lesions.(Wirth et.al.1996, Hughes et.al.1994)

Embolisation of PAVMs is advocated for all lesions with feeding vessels greater than 3mm. (White 1992, White et.al.1996). Rare complications of treatment include localised pleurisy, arrhythmia and rarely, paradoxical embolism of occluding device (White et.al. 1996, Haitjema et.al. 1996b, Haitjema et.al. 1995) . Early recurrence of an AVM can be due to failure to occlude all feeding vessels, late recurrence can occur following re-canalisation of a lesion (White et.al.1996, Sito et.al.1996). Small previously unidentified PAVMs may also increase in size and become significant during follow-up. Not all lesions are amenable to embolisation, some are still too large for embolisation and can be safely removed surgically. Treatment of patients with multiple small PAVMs remains a difficult problem (Wirth et.al.1996).

viii Complications of the Central Nervous System.

Some complications affecting the central nervous system such as embolic stroke and cerebral abscess result directly from loss of the pulmonary filtering capability due to direct arteriovenous communications. This has already been discussed.

Migraine occurs more commonly in hereditary haemorrhagic telangiectasia. (Steele et.al. 1993). The reason for this is unknown, possibly small assymptomatic cerebral arteriovenous malformations cause migrainous symptoms, or loss of pulmonary filtering capacity allows some vascular mediators to circulate in abnormal concentrations to the cerebral circulation, producing migrainous symptoms.

The frequency of cerebral arteriovenous malformations occurring in association with HHT is uncertain. Few studies have screened patients using cerebral CT scanning. An example of a characteristic high flow lesion is shown in **figure 1.2.**

Cerebral arteriovenous malformations (CAVMs) were found in 4/36 Dutch patients on CT scanning (Haitjema et.al.1995). Most lesions are assymptomatic, but catastrophic complications can develop without warning. Reports of CAVMs often show clustering within a family (Kikuchi et.al 1994, Lesser et.al. 1996, King et.al. 1977). This may be due to

genetic background influencing occurrence of CAVMs or specific gene mutations predisposing to such lesions as discussed later. Treatment of asymptomatic CAVMs is controversial, and whether or not to operate on a lesion depends on size and position of the lesion and estimated likelihood of bleeding.

ix Penetrance and Diagnostic Difficulties.

To date, the only assessment of penetrance of HHT is based on clinical criteria and incidence of unaffected obligate carriers of the disease. Three assymptomatic obligate carriers were seen in the French series of 103 patients (Plauchu et.al. 1989). When looking at penetrance in relation to age in this cohort, it was estimated it to be 75% by age 25 and 90% by age 45. In the British series, (Porteous et.al. 1992) 83% were symptomatic by age 26 and all by age 40.

Clinical diagnosis has always been based on the combinations of the presence of affected first degree relatives, frequent epistaxis, presence of a PAVM and classical telangiectases. Current opinion is to take the presence of two of these to denote affected status. However given the frequency of nosebleeds in the general population, Shovlin (Shovlin et.al.1994) recommended the proviso that frequent epistaxis with an affected relative should not be sufficient for a diagnosis in a research

setting. In a clinical setting, it may be appropriate to have a higher index of suspicion.

x. Pathological findings in Hereditary Haemorhagic Telangiectasia.

It is debatable whether the skin lesions in HHT are caused by an uncontrolled proliferation of endothelial cells, a developmental abnormality of blood vessels, or reduced resistance of the blood vessel to pressure, leading to dilatation and loss of normal capillary structure.

Three studies have looked at the microscopic structure of skin lesions. A study of a series of biopsies of skin lesions built up a picture of lesion development over time (Braverman and Jacobson 1992). Very early lesions showed dilatation of the post capillary venule with preservation of the capillary bed. As the lesions increased in size, the capillary segments disappeared, leaving a direct arteriovenous connection which enlarged under pressure. The lesions were associated with a perivascular cell infiltrate which consisted of a mixture of lymphocytes and monocytic cells.

Electron microscopy on a number of mucosal and cutaneous telangiectases showed that the lesions were dilated small venules with a thin endothelial cell wall (Hashimoto et.al. 1972). There was some evidence of weakness of junctions between endothelial cells. Often

dilatation resulted in cavernous spaces filled with fibrin and erythrocytes. The abnormal regions were not in association with smooth muscle cells or fibroblasts, implying that the primary pathology was one of endothelial cells.

A further study, (Menefee et.al 1975) confirmed the presence of dilated post-capillary venules, but noted that the basal lamina was often duplicated and thickened. This suggested that there was considerable endothelial cell death and regeneration. This might also be due to abnormal turnover of old basal membrane. Evidence of extravasation of blood was seen, including erythrocytes and deposition of fibrin, which could form the basis for new capillary sprouts, although no such sprouts were noted.

On balance, these studies suggest that HHT lesions develop as a result of progressive dilatation of post capillary venules, rather than as a result of endothelial proliferation. The triggering factor leading to the dilatation is unknown, and the reasons that only specific areas of skin and mucous membranes are affected are also uncertain. It has been suggested that thermal insult, ultraviolet light or a second genetic mutation may initiate dilatation of blood vessels.

C. Identification of Disease Genes

When there is no protein implicated in the causation of a genetic disease, the techniques used to identify the gene responsible are termed "reverse genetics". Typically the search for a gene takes place in four stages, initial linkage of the disease to a chromosomal location, narrowing the candidate interval within which the locus lies, identification of candidate genes within the candidate interval, and mutation analysis of candidate genes.

i. Linkage Analysis

Mendel's laws state that 2 loci segregate independently of each other, and at random. However, if two loci are close together on the same chromosome, they will co-segregate at a rate dependant on the distance between them. The greater the distance between the two loci, the greater the chance of a crossover event occurring at meiosis. The recombination fraction (denoted by the symbol θ) between two loci is therefore a measure of the distance between them. For two loci very close together such that there is never a recombination between them, $\theta=0$. For two loci so far apart that they segregate independently, $\theta=0.5$. Two loci are said to be linked when $\theta<0.5$, that is, when they segregate together more often than is expected by chance.

Linkage analysis is the method of statistical analysis for estimating θ, and determining whether the deviation of θ from 0.5 is statistically significant. The statistic used is the LOD score (denoted by Z). This value is a likelihood ratio derived using the formula:

$$Z(\theta) = \text{Log}_{10} \frac{L(\theta)}{L(\theta = 0.5)}$$

where the denominator is the likelihood of the marker data given that the recombination fraction is 0.5.

By convention, a LOD score of greater than 3 is taken as a significant score, demonstrating linkage. This approximates to a level of significance of 95% (Ott 1991). A LOD score of less than -2 is taken as exclusion of linkage.

Linkage analysis is commonly performed using the LINKAGE suite of programs (Terwilliger and Ott 1984). The MLINK program is used to calculate 2 point LOD scores for a range of recombination fractions between 2 loci.

For a rare autosomal dominant disease, it can be assumed that patients with the disease are heterozygous for a normal and a mutated allele. The

computer analysis also incorporates disease penetrance, usually estimated from clinical studies. Allele frequencies are also taken into account. Two point LOD scores are calculated between disease locus and highly polymorphic DNA markers. LOD scores derived from different families can be added together. However, if there is heterogeneity of disease locus, this may conceal a significant linkage result. Microsatellite markers are taken from published maps, typically that published by Genethon (Guapay et.al. 1994).

ii. Narrowing the Candidate Interval.

Once a chromosomal region is identified, further markers known to map to the region are analysed. Data from these markers can be formed into haplotypes within families. Critical recombination events within affected families, allow the identification of a smaller chromosomal region containing the gene.

iii. Identification of candidate genes and mutation analysis.

Once the smallest possible candidate interval is identified, candidate genes can be identified. Published genes or expressed sequence tags known to map to the region may then be assessed. Construction of a contig across the candidate region allows exact positioning of the candidate genes in relation to the microsatellite markers. Further genes can be identified from the clones of the contig using a variety of techniques

such as exon trapping and DNA direct selection (Parimoo et.al. 1991, Duyk et.al. 1990).

There are many ways of identifying mutations in candidate genes. If the gene is expressed at high enough levels in peripheral blood leucocytes, Reverse transcription followed by PCR (rtPCR) and sequencing allows rapid screening of larger genes. PCR amplification from genomic DNA (White et.al. 1989) allows use of a variety of mutation screening techniques. These include single stranded conformation polymorphism (Ainsworth et.al. 1990) , hydrolink$_{TM}$ gel electrophoresis (Keen et.al. 1991), denaturing gel electrophoresis, or chemical cleavage of mismatched bases (Cotton et.al. 1988). For smaller genes, sequencing may be practical (Sanger et.al. 1977). Genomic rearrangements can be sought by southern blotting following standard agarose gel electrophoresis or pulsed field gel electrophoresis. (Southern 1975, Schwartz and Cantor 1984).

D Identification of Endoglin as the HHT1 gene.

I. Linkage of HHT1 to Chromosome 9

In retrospect, the first evidence for linkage to chromosome 9 was in the Danish study (Vase and Grove 1986), which noted an association between HHT and blood group O in his population. Combined with the likelihood that the population studied from an isolated region of Denmark

has a high incidence of HHT due to a founder effect, this suggested that in this population, the mutation causing HHT is in linkage disequilibrium with the ABO locus. This was 8 years before linkage to this region was formally established.

Following a genome-wide search using polymorphic markers, linkage of HHT to chromosome 9 was established in a proportion of families. (McDonald et.al.1994, Shovlin et.al. 1994). More detailed analysis showed that the locus lay on 9q34 between markers D9S61 and D9S63, a 4cM interval.

ii.Identification of Mutations in the Endoglin Gene

Several genes lay close to, or within the candidate interval for HHT. As well as the COL5A1 gene and ZNF79, a recently identified protein, endoglin mapped to this region.

Endoglin is a TGF-β binding protein expressed predominantly in endothelial cells, and therefore a promising candidate for HHT. Further details of the role of Endoglin in cell signalling are given later.

The genomic structure of endoglin was elucidated by sequencing from a cosmid identified from a gridded genomic library. 3 mutations were identified in endoglin on screening a panel of 68 DNA samples from

unrelated HHT patients. (McAllister et.al. 1994a). The first three mutations identified, a premature stop, a 39 basepair insertion and a 2 bp deletion; all disrupted the endoglin molecule before the extracellular domain.

Five further mutations were identified on further screening of this panel. (McAllister et.al. 1995). None of these mutations occurred as a population polymorphism, and all would be expected to disrupt the endoglin molecule, confirming that mutations in the endoglin gene caused HHT.

Most of the British families with HHT did not show linkage to chromosome 9. Formal multipoint linkage analysis demonstrated exclusion of the 9q34 locus, indicating that there was a second unidentified locus for HHT. This was confirmed by other workers. (Heutink et.al. 1994, McAllister et.al 1994b, Porteous et.al. 1994, Shovlin et.al. 1994). This second unidentified locus was designated HHT2, and endoglin designated HHT1.

A subjective observation was made that the families which were linked to endoglin on 9q34 seemed to have a much higher incidence of pulmonary arteriovenous malformations reported than families for whom this locus was excluded.(Porteous et.al. 1994, McAllister et.al. 1994)

E. Initial identification of a second HHT locus on chromosome 12.

Initial linkage of 2 French families with HHT to chromosome 12 was reported (Vincent et.al. 1995). These families both came from the cohort of French patients previously published (Plauchu et.al. 1989), which had a low reported incidence of pulmonary arteriovenous malformations. This study defined a large candidate region between D12S345 and D12S368 of over 10cM, in the peri-centromeric region of chromosome 12.

F. Endoglin and its role in TGF-beta signalling.

An understanding of TGF-β signalling in the cell, and the role endoglin has in these pathways assisted in the identification of other molecules which are potential candidate genes for causing or modifying the HHT phenotype.

i.The TGF-beta receptor complex

The TGF-β receptor molecules are members of the serine-threonine kinase receptor family. These are a subgroup of receptor kinases that phosphorylate serine and threonine residues as a mechanism of downstream signalling. (ten Djike 1994a, Massague and Weis-Garcia 1996).

Initially the TGF-β receptors were classified by size. Two categories were identified, type I receptors of about 53kDa and type II receptors of 75kDa (Chiefetz et.al 1987). Later studies using both the TGF-β receptors and other family members have built up a model of the interaction between different receptor types. Association of TGF-β with the TGF-β type II receptor causes formation of a complex between ligand, type II receptor and type I receptor. The type II receptor phosphorylates the type I receptor, activating it. The type I receptor is then capable of downstream signalling alone (Ebner et.al. 1993, Franzen et.al. 1993, Attisano et.al. 1993, Attisano and Wrana 1996, Massague and Weis-Garcia 1996).

The type I receptor is therefore characterised by an inability to bind ligand alone, but it is essential for downstream signalling. All type I receptors identified to date have a specific intracellular region, the GS domain with a key amino acid sequence, TTSGSGSG. (Franzen et.al. 1995) This is the site of phosphorylation for activation. Mutations of this site reduce downstream signalling. A specific mutation downstream from this site can cause constitutive activation of the receptor molecule. (Attisano et.al 1996).

It is not known whether the type I receptor, once activated is inactivated by dephosphorylation, dissociation or internalisation and degradation.

ii.Endoglin and Betaglycan - "type III" receptors.

Two molecules have been identified which bind TGF-β and are not classical serine threonine kinase receptors in the model discussed above. Betaglycan, is a proteoglycan which can either be membrane bound or released in a soluble form. The intracellular domain is small and has no discernible signalling motifs. (Lopez-Casillas et.al. 1991). When co-expressed with TGF-β receptors type I and II, it increases the binding affinity for TGF-β2. Cells lacking betaglycan but expressing TGF-β receptors type I and II are insensitive to TGF-β2, but can be sensitised to it by expression of betaglycan. (Lopez-Casillas et.al 1993). Soluble betaglycan can sequester TGF-β and in vivo may act as a reservoir in the extracellular matrix. (Lopez-Casillas et.al. 1994).

Endoglin is related to betaglycan, with considerable homology of the extracellular domain and a short intracellular domain. It was initially identified in a pre-B leukaemic cell line and later shown to be expressed in endothelial cells, activated monocytes and placenta. Studies in mice have also shown expression in the stromal cells in connective tissue of many different organs. (O'Connel et.al. 1992, Cheifetz et.al. 1992, St.-Jacques et.al.1994a,b, Lastres et.al.1992, 1996). It can be found in a soluble form in blood, presumably after release from the endothelial cell surface (Lopez-Casillas et.al 1994). It has a high affinity for TGF-β1 and 3, but not TGF-β2. Studies suggest formation of a signalling complex between TGF-β

receptors and Endoglin. (Yamashita et.al. 1994a,b, Lastres et.al. 1996). The true effect of endoglin on signalling *in-vivo* remains to be determined.

iii TGF-β and related Cytokines.

TGF-β belongs to a group of related cytokines that are the ligands for the serine-threonine kinase receptor family. This diverse family of intercellular signalling molecules includes the different isoforms of TGF-β, as well as the bone morphogenic proteins, inhibins, mullerian inhibitory substance and glial cell derived neurotrophic factor. Each signalling molecule is made up of a heterodimer or homo-dimer. The monomeric subunit of all signalling molecules in this group each have an internal structure of 6 cysteines forming 3 disulphide bonds and a seventh cysteine that forms a bond with another monomer. (Massague and Weiss-Garcia 1996,). The monomer is synthesised as a larger precursor, with activation by removal of the N-terminal portion of the molecule after dimerisation.

There are 3 different potential subunits of TGF-β, which usually form homodimers to give the TGF-β isoforms 1,2 and 3. Affinity of receptors for these molecules varies. the TGF-β receptor complex has a lower affinity for TGF-β 2, however its affinity is increased in the presence of betaglycan (Cheifetz et.al. 1990, Lopez-Casillas et.al. 1994). Endoglin has higher affinity for TGF-β 1 and 3 than TGF-β 2 (Yamashita et.al. 1994).

The different function of these molecules in vivo is highlighted by the different phenotypes shown in knockout mice. The TGF-β 3 knockout mouse shows cleft palate and pulmonary hypoplasia (Kaartinen et.al. 1995, Proetzel et.al. 1995), whereas, the TGF-β 1 knockout mouse shows a chronic inflammatory disease or defects in yolk sac formation depending on genetic background (Dickson et.al. 1995, Martin et.al. 1995). A putative modifier locus has been mapped in the mouse (Bonyadi et.al. 1997), located on mouse chromosome 5.

iv.Intracellular Signalling.

Advances have been made very recently into understanding the mechanism for intracellular signalling of serine-threonine kinase receptors.

The specificity of downstream signalling of a serine-threonine kinase receptor complex appears to be determined by the kinase domain of the type I receptor (Carcamo et.al. 1994, Attisano et.al 1993, Massague and Weiss-Garcia 1996). The TGF-β receptor type I and Activin IB receptor both signal via the same pathway in mink lung epithelial cells but require their specific type II receptor and ligand for the signalling to take place. Their kinase domains are 90% identical at the amino acid level (Carcamo et.al. 1994).

Studies to look for proteins interacting with the intracellular domain of the TGF-β receptor type I have shown several possible candidates for downstream signalling. These include Farnesyl Transferase (Kawabata et.al. 1995) and FKBP12 (Wang et.al 1994), both of which were identified in a yeast two hybrid assay.

Recent attention has focused on the family of MAD (Mothers against decapentaplegic) proteins. MAD originally identified as a protein involved in signalling of DPP (decapentaplegic) in Drosophila (Niehrs 1996, Liu et.al. 1996, Graf et.al. 1996, Hoodless et.al. 1996) which was then shown to activate transcription of a reporter construct (Liu et.al. 1996). The human homologue, MADR1 has been shown to move from the nucleus in response to BMP2 (Bone morphogenic protein 2) signalling in conjunction with the BMP type I and II receptors (Hoodless et.al. 1996). An entire family of MAD proteins was subsequently identified (Riggins et.al.1996). Five new human members of this group were identified by their homology to drosophila MAD (Riggins et.al. 1996). Previously identified members include the DPC4 (deleted in pancreatic carcinoma) gene, which shows less homology to the other known MAD proteins.

All MAD proteins identified to date have a two highly conserved domains, the N- terminus MH1 domain and the C-terminus MH2 domain. The region between is highly variable. Two serines at the C-terminus are thought to be

the site for phosphorylation of the molecule when it interacts with the type-I receptor. (Attisano personal communication). Phosphorylation of the molecule leads to nuclear localisation (Hoodless et.al. 1996, Chen et.al. 1996, Liu et.al. 1996) However, DPC4 lacks the phosphorylation site, but can localise to the nucleus only in presence of activated MADR1. (Zhang et.al. 1996). It is thought, therefore to be a component of the same signalling pathway.

v.Other serine-threonine kinase receptors.

There are many other members of this receptor family that have been identified in different species. In drosophila, punt is a type II serine threonine kinase receptor (Letsou et.al. 1995), which in combination with either Sax or Tkv type I receptors, responds to DPP (and human bone morphogenic protein 2) as a ligand.

In humans, the activin type Ia and Ib receptors signal in combination with the activin type II receptor and activin (Attisano et.al. 1996). They signal through different pathways, the type Ib receptor using the same pathway as the TGF-β type I receptor (Carcamo et.al. 1994). In humans, bone morphogenic protein receptors and the mullerian inhibitory substance receptor have been identified (Baarends et.al. 1994, Massague and Weiss-Garcia 1996). A further receptor, the Activin receptor like kinase I (ALK-1) has been identified (Attisano et.al. 1993, Ten Djike et.al. 1993).

This type I receptor is expressed predominantly in endothelial cells, demonstrated by rtPCR from a range of cell lines (Attisano et.al. 1993). Neither its true ligand in vivo nor its downstream signalling mechanism are known.

The large number of serine threonine kinase receptors and ligands make it possible for both inter-related and completely separate serine threonine kinase pathways to exist within the same cell. The combination of specific type I and type II receptors expressed by a particular cell decide if and how it responds to molecular signals from surrounding cells.

vi.The role of TGF-β in the endothelial cell.

Several studies have looked at the effect of TGF-β1 on the growth of endothelial cells in culture. Microvascular endothelial cell monolayers in 2 dimensional culture grow rapidly, but addition of TGF-β1 inhibits this and induces differentiation with expression of alpha-smooth muscle actin, PDGF α and PDGF β receptors and fibronectin (Mandriota et.al. 1996, Sankar et.al. 1996, Folkmann and D'Amore 1996, Lastres et.al. 1996).

When the cells are growing in a 3 dimensional matrix, the addition of TGF-β induces the formation of tubules by the cells, but does not affect the growth rate. It still induces fibronectin production by the cell. (Madri et.al 1988, Merwin et.al 1991, Merwin et.al 1990, Sankar et.al 1996). TGF-β2

can have the same effect on cell proliferation but requires the expression of betaglycan for the cells to be sensitive. (Sankar et.al 1996).

The TGF-β type II receptor is expressed at lower concentration in cells which are growing in three dimensional culture, suggesting that the concentration of this receptor may be important in controlling cell responsiveness (Sankar et.al. 1996). Potentially this receptor may interact with other type I serine threonine kinase receptors as well as the TGF-β type I receptor.

Studies in vivo have shown that TGF-β1 can cause increased vascular proliferation and fibrosis, but can also accelerate wound healing. (Roberts et.al.1986, Mustoe et.al 1987).

2. Materials and Methods

A. Aims of Investigation

To identify the HHT2 gene, with comparison of clinical features between HHT1 and HHT2.

B. Plan of Investigation

1. Assessment of new families with HHT.
2. Analysis of clinical data from HHT families.
3. Initial linkage studies to identify the locus involved in each family.
4. Haplotype analysis of families to define the candidate region for HHT2.
5. Study of known genes and ESTs in this region as candidates for HHT2.
6. Determination of the genomic structure of the HHT2 gene.
7. Further mutation analysis of the HHT2 gene.

C. Ascertainment of Families with HHT

Samples and data were already available from British and American families that had previously been assessed and published. (Porteous et.al. 1994, McAllister et.al. 1994a, Heutink et.al. 1994). Further British families were identified through family members that had contacted the Clinical Genetics Service in Edinburgh. 2 large American families contacted the department of Genetics at Duke University, wishing to participate in HHT research.

Initial diagnostic criteria used for a definite diagnosis of HHT were the presence of two of the following:

1. Affected first degree relative
2. More than monthly nosebleeds
3. A classical telangiectasis
4. Presence of a known Pulmonary Arteriovenous Malformation

Subsequent experience as described in the text led to modification of these criteria, individuals with an affected relative and nosebleeds alone were classified as of uncertain diagnosis during haplotype analysis. This was due to a perceived high frequency of unnaffected individuals who had monthly nosebleeds.

All families were assessed during large family gatherings. Individuals were asked to fill in a questionnaire, before being briefly interviewed to clarify details. The questionnaire was designed to record details necessary for diagnosis of HHT as well as to gather clinical information on presence, frequency and severity of the reported complications. These included nosebleeds, gastro-intestinal haemorrhage, hepatic complications, arteriovenous malformations and migraine. A data recording sheet was also used to note and summarise clinical findings. The questionnaire and recording sheet are given in **appendix C**.

After completion of details in the questionnaire, the individual was examined for telangiectases on the hands, face and mucous membranes.

Pulse oximetry was performed to screen for pulmonary arteriovenous malformations. The oxygen saturation was measured for 5 minutes when supine and then 5 minutes standing. A drop of saturation of 2% was taken as possibly indicative of a PAVM, requiring further investigation. Unfortunately, when seen in American family members it was not possible to pursue this, except by detailed history. 2 British family members were further investigated by the Department of Respiratory Medicine at the Hammersmith hospital in London. Neither individual had a PAVM.

A blood sample of 10mls was anticoagulated with potassium EDTA for later DNA extraction. Blood samples were either stored at -20 Celsius until DNA was extracted, or DNA was extracted within 24 hours. DNA extraction was performed using the protocol given below.

D. Analysis of Polymorphic Markers and Linkage Analysis

Polymorphic markers chosen from the regions of interest were amplified from patient DNA using the polymerase chain reaction and analysed by polyacrylamide gel electrophoresis using the protocol given. Allele sizes were either arbitrarily assigned or estimated using the standard CEPH individual 134702.

2 point linkage analysis of the British families was initially performed using MLINK from the LINKAGE version 4.9 suite of programmes running on IBM compatible computers (Terwilliger and Ott 1994). A penetrance of 90% was assumed for HHT2 gene carriers, with equal allele frequencies. Multipoint analysis for family 5 was performed using the LINKMAP program.

Later analysis of the British families and of the American families was also performed using MLINK, with the following age related penetrance profile: age 0-20, penetrance 0.24, age 21-40 penetrance 0.73, age >40

penetrance 0.97. Allele frequencies for each marker were taken from published data.

Haplotype analysis was performed manually adding allele data for each marker in order of position to a printed family tree. Sites of definite recombination were noted. For the families described in this study, no inconsistent haplotypes were found to suggest non paternity or sample error.

The smallest candidate interval defined by haplotype analysis was compared to the physical map supplied by R.Kucherlapati. This gave more precise details of the size of candidate interval, marker position and genes contained within it.

E.Comparison of Clinical Data.

Clinical data between families was inspected, both as raw data, and in the form of a two by two table. Typically the two by two table was used to compare the presence or absence of a feature with linkage or exclusion of HHT1 on chromosome 9. Statistical analysis was performed using the Fishers exact test or the χ^2 test depending on sample size. Results were calculated using the MINITAB or SPSS. packages.

F.Characterisation of ESTs in the Candidate Region.

Phage clones from the endothelial cell cDNA libraries EC1 and EC2 (Ginsburg et.al. 1985) corresponding to ESTs were isolated using standard techniques. After 3 rounds of plaque purification, the single clones identified were PCR amplified and sequenced using ^{33}P labelled dideoxy termination reactions as described below. Sequence identified was entered into a computer file using the MAC-vector package. Sequence was inspected using this software for overlap, open reading frames, and restriction sites. Homology to known proteins was sought using the BLAST program to search the GENBANK non-redundant nucleotide and non-redundant protein databases.

G.Identification of a PAC clone containing the ALK-1 gene.

DNA pools from a 3 dimensionally arrayed PAC library (described in Iannou et.al 1994) were screened by PCR using the primers

5′ CGCGTGTCACACTTCATGGCTC 3′ and

5′ ATCAGAAGGCCTTTCCTGGGGG 3′

which are known to specifically amplify a fragment of the 5′ untranslated region of the ALK-1 gene. (Attisano et.al. 1993)

This reaction identified the co-ordinates 5X3 5Y1 5Z1 in the grid, indicating a 384 well plate number 265. An imprint of this plate was stamped into agarose, the colonies were grown and lifted onto Hybond N membrane,

before being probed using the ALK-1 cDNA.(Kindly supplied by Attisano) This identified a single colony in position C-14. This clone 265_C_14 (renamed GA-1) was digested using a range of common restriction endonucleases, resolved by agarose gel electrophoresis and southern blotted to hybond-N membrane. This was hybridised to radiolabelled ALK-1 cDNA as described below.

H.Subcloning of GA-1

A Bgl II digest of this subclone was purified using phenol-chloroform extraction and ligated into the BamHI site of the stratagene pBluescript phagemid. Clones which hybridised to the ALK-1 cDNA were isolated and characterised by digestion with EcoRI and Not1.

These digests were also resolved on southern blots and hybridised to specific primers complementary to the ALK-1 cDNA sequence, to identify which fragment of the gene was contained in each clone.

I.Identification of Intron-Exon Borders in the ALK-1 gene.

Intron exon borders within the ALK-1 gene were identified by PCR using the PAC clone GA-1 as template, and primers from the cDNA sequence of ALK-1. The sequence of primers used for this analysis is given in **table 2.1.** Wherever a PCR fragment generated was larger than the fragment expected from the cDNA sequence, the fragment was assumed to contain

an intron. Such fragments were sequenced as described. The sequence generated was compared to known cDNA sequence, and the intron/exon border was identified where the derived sequence diverged from that of the cDNA. For some of the intron/exon borders more sequence was generated directly from the subclones to clarify sequence of regions not clearly sequenced from PCR products.

J.Mutation analysis of patient samples.

Primers were designed to amplify each of the coding exons of ALK-1 (exons 2 to 10). The sequence of these primers is given in **table 2.2**. The amplification conditions were optimised to give a clear single product band, using either buffer supplied by Boehringer-Mannheim, or that described (Gyapay et.al. 1994). In all cases, the magnesium concentration used was 1.5mM.

The single product band was isolated on low melting point agarose gel electrophoresis and extracted by freezing followed by compression in a micro-centrifuge. The supernatant was sequenced as described below. For each termination reaction, samples for all 12 individuals were loaded adjacent to each other on the sequencing gel, making identification of sequence variations easier. One set of 4 sequencing reactions was loaded in the conventional order ACGT as a ladder on each gel.

Primer ID	Primer sequence	Exonic Location and orientation
561	TGGCTCTTACTCCACCTCTCTTGCTCCT	Forward
124	GAGAGTCCAGTCTCATCCTGAAAGC	Forward
489	TGTGGCACGGTGAGAGTGTGGCC	Forward
501	GCTCTGAGGCTAGCTGTGTC	Forward
505	GTGGACTGACATCTGGGCCTTT	Forward
507	AGAAGGTGGTGTGTGTGGATCA	Forward
460	CCTCCCCTCCTCCCGCAC	Reverse
280	TGAGCCACTCCCTGTGGTGCAGTC	Reverse
490	CCCGGAACCAGGACTGTTCATCCCTC	Reverse
125	ACACTGCAGGTTGCTCTTGACC	Reverse
504	GTAGGACTCAAAGCAGTCCGTGC	Reverse
506	GTCCTCAAAGCTGGGGTCATTG	Reverse
560	CTGCAGGCAGAAAGGAATCAGGTGCT	Reverse
JB3	AGGTGGCCCCGGTCCGCCGAAGGCTA	Forward
JB4	CTGGGTCACCAAGGCCATCAGCAGCATCAGAA	Reverse

Figure 2.1 Sequence of key primers used for the identification of intron-exon borders in the ALK-1 gene.

Exon	5' Assay primer	3' Assay primer	Annealing Temp.
II	CTCTGTGATTTCCTCTGGGCA	TACATTCTCCCCAGCTTCTCAA	62
III	AGCTGGGACCACAGTGGCTGA	GGAGGCAGGGGCCAAGAAGAT	64
IV	AGCTGACCTAGTGGAAGCTGA	CTGATTCTGCAGTTCCTATCTG	60
V	AGGAGCTTGCAGTGACCCAGCA	ATGAGAGCCCTTGGTCCTCATCCA	68
VI	AGGCAGCGCAGCATCAAGAT	AAACTTGAGCCCTGAGTGCAG	60
VII	TGACGACTCCAGCCTCCCTTAG	CAAGCTCCGCCCACCTGTGAA	65
VIII	AGGTTTGGGAGAGGGGCAGGAGT	GGCTCCACAGGCTGATTCCCCTT	65
IX	TCCTCTGGGTGGTATTGGGCCTC	CAGAAATCCCAGCCGTGAGCCAC	68
X	TCTCCTCTGCACCTCTCTCCCAA	CTGCAGGCAGAAAGGAATCAGGTGCT	65

Table 2.2 Primers used to amplify each exon of the ALK-1 gene for mutation analysis. The annealing temperature of the PCR reaction used with each primer set is given.

Sequence variations that were identified were entered into MACvector and the effect of the sequence variation calculated. Where possible, a change of predicted enzyme digest pattern was identified, for use as a polymorphism screen.

For each putative disease causing mutation identified, a panel of over 100 individuals from the normal population was screened to check that the mutation did not arise as a common population polymorphism. The samples used for this came from married-in members of the families collected and control samples available locally. In all but one mutation, 112 individuals were screened, estimated to give a 90% chance of identifying a population polymorphism. The remaining mutation was checked in a panel of 106 individuals.

The mutated region was also compared to related proteins in GENBANK using the BLAST program, to estimate the degree of homology present in the region that was mutated.

K.Degenerate PCR to identify MAD protein homologues

To determine the sequence of degenerate primers calculated to amplify MAD protein homologues, sequence of known MAD protein genes was aligned manually. Conserved motifs from the MH1 and MH2 domains were used to design forward and reverse primers respectively. PCR conditions

were varied as described in the protocol. Template used was either cDNA prepared from endothelial cell lines, or DNA extracted from an aliquot of the EC-1 or EC-2 libraries. Product of interest was cloned using the T/A cloning system (Invitrogen), and sequenced using the usual protocol.

L. Protocols

i. DNA extraction

(Adapted from local laboratory protocol)

DNA extraction was performed according to the protocol below for all samples originating from the U.K. All American samples were extracted using the Puregene DNA extraction kit (Gentra Systems) with the protocol supplied.

10mls of blood is mixed with 30mls of red cell lysis buffer in a closed 50ml tube. This was placed on ice for 30 minutes with intermittent inversion of the tube. When the mixture had become clear, indicating red cell lysis, the tube was centrifuged at 3000g for 10 minutes, and the supernatant discarded. The white cell pellet was resuspended in 5ml of red cell lysis buffer, and re-centrifuged at 3000g, with discarding of the supernatant. this step was repeated if necessary until the pellet was white.

The pellet was resuspended in 3ml of nuclear lysis buffer,200μl of 10%SDS and 300μl of 10mg/ml proteinase K. The tube is incubated overnight at 37°C. After 16hours, 1ml of 6M NaCl was added, with vigorous agitation for 15 seconds. The tube was then centrifuged at 3000rpm for 20 minutes at 4°C. The supernatant was transferred to a new container, and 2

volumes of ethanol added, before being inverted several times. The resulting DNA precipitate was spooled out and resuspended in 250 to 500μl of TE solution. This was mixed for at least 16 hours on a rotating mixer to allow resuspension of the DNA.

ii. Restriction endonuclease digestion.

Restriction endonuclease digestion was performed using the same protocol for all types of DNA substrate. Using 1X concentration of the recommended buffer for an enzyme, 1 to 10 units of enzyme were used per 1μg of DNA. Where double digest was performed, a buffer was chosen that permitted digest by both enzymes, according to manufacturers recommendations.

iii. Phenol-chloroform DNA extraction / DNA precipitation

(Adapted from Sambrook et.al. 1989)

The DNA sample to be purified is mixed with an equal volume of phenol/chloroform solution, and mixed vigorously. The aqueous layer is removed and mixed with an equal volume of chloroform/isoamyl alchohol. The aqueous layer is removed once more. A tenth volume of 3M sodium acetate was added and 2 volumes of ethanol. The mixture was frozen at -70°C and then centrifuged at 14000g for 20 minutes. The supernatant was removed and the pellet washed in 70% ethanol before being air dried and resuspended in water or TE solution pH 8.0.

iv. Polymerase chain reaction (PCR)

All PCR conditions were variations on single basic protocol:

Forward primer:	100ng
Reverse primer	100ng
TAQ DNA polymerase	1 unit
Buffer	1X
Template	10ng to 500ng
Deoxy-ribonucleotide triphosphates	0.2mM of each
Final volume	50μl

The buffer used was either that supplied by Boehringer-Mannheim or that use by Gyapay et.al. (1994), referred to in this text as "Weissenbach" buffer, which gave a final concentration of 10mM Tris pH9, 50mM KCl, 1.5mM $MgCl_2$, 0.1% Triton x-100 and 0.01% gelatin.

All reactions were overlayed with mineral oil and thermocycling was performed using a range of purpose built machines for single tubes or 96-well plates. All reactions started with an initial denaturing temperature of 95°C for 5 minutes. Then typically 95°C for 45 seconds, annealing temperature for 1 minute and 72° C for a variable length of time. 35 to 40 cycles were used. A terminal extension step of 10 minutes was used in most reactions.

Variables in the PCR reaction

*Hot Start:*In most cases the fidelity of the PCR reaction was improved by the addition of the TAQ polymerase during the initial 95°C denaturing step.

This reduces non-specific product in the initial reaction steps, as the primers do not have the opportunity to anneal at a temperature below the annealing temperature chosen for the PCR reaction. It also reduces primer-dimer formation by a similar mechanism.

Extension time: The extension time was determined by the length of the anticipated product. For fragment less that 200bp 30 seconds was used. Otherwise, 1 minute per kb length of product was used. For exon to exon PCR reactions, extension times of 5 to 7 minutes were used for initial trials of primer pairs.

Annealing temperature: At higher annealing temperatures, primer annealing is more specific, however product yield is usually reduced. The annealing temperature was usually set at 2°C below the TM (melting temperature) of the primer with the lowest TM. This was estimated by allowing 4°C for each C or G in the sequence and 2°C for each A or T. When a reaction failed to work, a lower annealing temperature was tried. If there were additional undesirable products in the reaction, the annealing temperature was raised by 2°C to attempt to remedy this.

Primer design: Most primers were designed manually. The ideal sequence was felt to be 50-60% GC content with a TM of around 65°C. Short repeated stretches in the primer were avoided where possible. There was

a subjective impression that primers ending in an A or a T gave better results.

Touchdown PCR programmes: To increase product specificity, touchdown PCR starts with a high annealing temperature, producing small amounts of specific product, increasing the concentration of correct template. A typical touchdown programme started with an annealing temperature 6° above that normally used for 3 cycles, 4° above for 3 cycles and then 2°above for 3 cycles, followed by 30 cycles at the normal annealing temperature.

Magnesium concentration: For all reactions, a constant magnesium concentration of 1.5mM was used. Increasing the concentration to 2.0 or 2.5mM would have been expected to increase product concentration but reduce specificity of the product formed.

Degenerate PCR reactions: using degenerate PCR primers, which are a mixture of 500 to 1000 different primer sequences causes a range of difficulties, some of which can be compensated for by altering reaction conditions. Primer concentrations were varied between 10 and 100 fold that usually used. To increase reaction specificity, hot start was used in all cases. All successful reactions used a touchdown sequence to increase specific product while reducing background. A range of annealing

temperatures were used, based on the average TM of the primers designed.

Exon to Exon PCR reactions: For the exon to exon PCR reactions, many different primer combinations were used. The TMs of the different primers were often very widely spaced. The length of product was unknown. Reactions all used a hot start with an extension step of 5 to 7 minutes. Using PAC DNA as a template this proved effective.

v. Agarose gel electrophoresis

(Adapted from Sambrook et.al. 1989).

Agarose gel electrophoresis was used for the size separation of DNA fragments between the sizes of 100bp and 15kb. The size of expected fragment determined the gel concentration used; 0.7% for the largest fragments, 1% for fragments up to 1kb and 1.5 to 2.5% for smaller fragments. Gels were made with 0.5X TBE containing 0.3μg per ml of ethidium bromide.

Gels were run in 0.5X TBE running buffer, and photographed with UV transillumination to visualise the ethidium stained DNA.

Fragment sizes were calculated by comparison to standard ladder (Gibco BRL 1kb ladder, or Gibco BRL 100bp ladder). Samples were mixed with loading dye, before being loaded on the gel.

vi. Southern blotting

(Adapted from Sambrook et.al. 1989, Southern 1975)

In all cases, DNA was transferred from agarose gels to Hybond-N membrane (Amersham) using the same protocol. The gel was photographed under UV transillumination, adjacent to a ruler to allow sizing of fragments. The gel was then inverted onto a standard capillary blotting apparatus. The blotting solution used was 0.4M NaOH in all cases. A piece of Hybond-N membrane was placed on the inverted gel, and a layer of 3MM paper followed by paper towels on top of this. Blotting was allowed to take place for at least 12 hours.

The membrane was then removed and washed in 2XSSC. It was fixed by exposure to UV light (1200J on the automatic setting of a stratagene stratalinker).

vii. Pulsed Field gel electrophoresis

(From pharmacia gene navigator protocol)

This was used for the size separation of fragments between 10kb and 100kb in size. Running conditions for a pharmacia gene navigator were

calculated empirically. A 1.2% agarose gel, with 0.15X TBE running buffer was used. A gene navigator (pharmacia) electrophoresis apparatus was used with a hex electrode. Gels were run for 6 hours with a switching time starting at 0.4 seconds increasing by 0.1 seconds per hour.

Gels were stained in ethidium bromide solution (0.1μg per ml), photographed and southern blotted as described above. A depurination step was not found to be necessary.

viii ^{32}P labelling of DNA probes

Random oligo-labelling. Used for all sizeable DNA fragments. For some probes, random hexamer primed labelling using the protocol given by Feinberg and Vogelstein (1984) was used. For others, the prime-it kit (stratagene) which uses random nonamers was used. α-^{32}P dCTP was used as the incorporated label. Labelled probe was heated to 100°C with 1000μg of salmon sperm DNA before addition to the hybridisation mix.

End labelling: Oligonucleotides, when used as probes were labelled using γ–^{32}P ATP and T4 kinase (gibco-BRL) according to the following protocol:

γ–^{32}P ATP 10μCi/μl	5 μl
100ng Oligonucleotide + water	6 μl
5X forward reaction buffer	3 μl
T4 Kinase	1 μl

Probe was heated to 100°C before addition to the hybridisation mixture.

ix.Hybridisation

Prehybridisation of the membrane was performed at 65°C with 15 to 25ml of hybridisation mixture (depending on size of hybridisation jar) for at least 1hour. Probe was boiled as described above and added directly to the hybridisation jar. Random hexamer and nonamer labelled probes were hybridised for 4 to 24 hours at 65°C. Oligonucleotide probes were hybridised for 4 to 24 hours at 48°C.

The membrane was then washed several times for 10 minutes at room temperature, with further, more stringent washes at 48°C and 65°C if required. Autoradiography was performed with intensifying screens at -70°C for 4 to 48 hours.

ix. Purification of PCR product.

PCR products were loaded into a 1% low melting point agarose gel prepared with ethidium and 0.5X TBE as described above. Electrophoresis was carried out for an appropriate length of time to separate the required fragment size from other bands. This band was then excised under UV transillumination, and wrapped in parafilm. This package was frozen at -70°C for 20 minutes. Manual pressure was then applied to the gel slice at room temperature, and the resulting solution pipetted into a fresh tube.

Where DNA quality was particularly important, notably for ligation reactions, the Qiaex II (Qiagen) kit was used as an alternative method.

x. Sequencing

All sequencing was performed using the thermosequenase cycle sequencing kit (Amersham). This uses a variation of the dideoxy termination method of Sanger(Sanger et.al. 1977), with ^{33}P radiolabelled dideoxy terminators. This protocol proved effective across a wide range of template concentrations.

All PCR products were gel purified as described above before sequencing. For plasmid clones, 100ng to 1000ng of DNA template from any form of DNA preparation (as described below) was used.

xi. Sequencing gel electrophoresis

Sequencing reactions were resolved by electrophoresis through a 7M urea/6% poly-acrylamide sequencing gel using 1X TBE as the running buffer. Run times were between 2 and 6 hours, with a constant power of 65W. Samples were denatured at 95°C with formamide loading buffer provided by Amersham, before loading on the gel.

Gels were fixed with 10% methanol / 10% acetic acid, transferred to 3MM paper and dried. Autoradiography was carried out without intensifying screens for 12 to 48 hours.

For microsattelite markers, 25μl of sample was mixed with 6 μl of formamide loading dye (95% formamide, 10mM NaOH, 0.005%bromophenol blue, 0.005% xylene cyanol). This was heated to 95°C before loading 3 to 7μl on a gel. Run times were between 90 minutes and 4 hours.

xiv. Bacterial colony isolation

(Adapted from Sambrook et.al. 1989)

BamHI digested pBluescript vector (stratagene) was treated with alkaline phosphatase (Gibco BRL) before phenol chloroform extraction. The ligation reaction contained 10ng of digested vector and approximately 80ng of insert in a final volume of 20μl. 1 unit of T4 DNA ligase was used (Gibco BRL). The reaction was allowed to take place at 12°C overnight.

3μl of ligation reaction was used to transform DH5α competent cells (Gibco-BRL) by heat shock, according to the protocol supplied. 10% and 40% of each transformation was plated onto LB agar plates containing 40μg per ml of X-gal and 100μg per ml of ampicillin. These plates were incubated overnight at 37°C. Single white colonies were picked with a wire loop and transferred to a second LB/ampicillin plate in a 10 by 8 grid.

After an overnight incubation, an impression of the grid was transferred to a circle of hybond-N membrane. This was placed on 3MM paper soaked in 0.4M NaOH, 1.5M NaCl for 2 minutes, and then transferred to 3MM paper soaked in Tris HCL pH 7.5. for 2 minutes and then 2XSSC for 2 minutes. The filter was then air dried, before being UV fixed using a stratalinker (stratagene) on the auto-crosslink setting. The membrane was then probed by hybridisation to radiolabelled ALK-1 cDNA probe. The resulting autoradiogram was aligned to the original agar plate with gridded colonies and positive clones picked for culture.

Single clones were grown with constant agitation overnight at 37°C in 2ml of LB medium with 100μg per ml of ampicillin. DNA was prepared using either Qiaprep kits (Qiagen) or the modification of the boil-lysis preparation given below:

2ml of bacterial culture was centrifuged at 14000Xg for 5 minutes and supernatant discarded. The cell pellet was resuspended in 50μl of boil lysis solution, agitated at room temperature for 5 minutes, heated to 100°C for 1 minute and placed on ice. The resulting solution was centrifuged at 14000xg for 15 minutes, and the DNA contained in the supernatant removed to a separate tube for digestion or other processing.

xv. Reverse transcription

Reverse transcription was carried out according to the protocol supplied with superscript II reverse transcriptase (Gibco-BRL) 1st strand synthesis was primed using random oligomer primers or $(T)_{18}$. 2μg of total endothelial cell RNA and 500ng of primer were used in each reaction.

xvi Phage plaque purification

This was carried out according to the protocol provided in Sambrook et.al. 1989.

M. Solutions and Buffers.

Red cell lysis buffer

155mM NH_4Cl
10mM $KHCO_3$
0.1mM EDTA
(41.45g NH_4Cl, 5g $KHCO_3$, 1.85g EDTA pH adjusted to 7.4 with Conc HCl, to a final volume of 5l in distilled water. Solution filtered and autoclaved)

Nuclear Lysis Buffer

10mM Tris HCl
0.4M NaCl
20mM EDTA
(2.4gTris HCl, 46.8g NaCl, 1.5g EDTA, to a final volume of 2l in water. Filtered and autoclaved)

TE pH 8.0

10mM Tris HCl
1mM EDTA pH 8.0
(1.21g Tris HCl, 0.37g disodium EDTA, pH adjusted to 8.0 with HCl, to a final volume of 1l in water. Solution filtered and autoclaved.)

Phenol/Chloroform
25:24:1 mixture of buffered phenol, chloroform and isoamyl alchohol

Chloroform/Isoamyl Alchohol
24:1 mixture of chloroform and isoamyl alchohol

6X sample loading dye

15% ficoll
0.001% xylene cyanol
0.001% orange G

10X TBE

0.9M Tris/borate
37.2g EDTA

(540g Tris base, 275g Boric acid, 37.2g EDTA made up to 5l in water)

Hybridisation solution

Bovine serum albumin (BSA)	0.5g
Polyvinyl pyrollidine	0.5g
Ficoll	0.5g
SDS	1g
Sodium pyrophosphate	1g
20X SSC	250ml
Water	744ml

Wash Solution

4X SSC
0.1%SDS
0.1% Sodium Pyrophosphate

Boil Lysis Solution

10mM Tris-HCl pH8
1mM EDTA
15% w/v Sucrose
2mg/ml lysozyme
0.2mg/ml RNAse
0.1mg/ml Bovine Serum Albumin

Other Solutions.

All other solutions used were made according to the methods given in

Sambrook et.al. 1989.

Chapter 3. Clinical Ascertainment and Linkage Analysis

A. Assessment of New Families with HHT.

In addition to 4 large British families which had already been investigated and shown not to be linked to chromosome 9q34 (Porteous et.al. 1994), five new families were identified, 3 in the United Kingdom and 2 in the USA. The pedigrees of these families, numbered 5,6,7,8 and 9 (USA) are shown in **figure 3.1**.

Individual data was recorded on the questionnaire and data sheets given in **appendix C**. All individuals were screened for pulmonary arteriovenous malformations by pulse oximetry as described. The spectrum of features within each family is given in tabular form in **appendix B**.

Table 3.1 summarises the proportion of individuals in each family with each recognised complication. It is notable that in families 6 and 8, the incidence of pulmonary arteriovenous malformations was 2/5 and 2/8 respectively, while in families 5,7 and 9, there was no history of PAVM or evidence of such a lesion on oximetry.

The total incidence of hepatic involvement is low at 2/45, both these individuals being asymptomatic. The incidence of gastro-intestinal haemorrhage was 7/45, a similar proportion to that found by other workers.

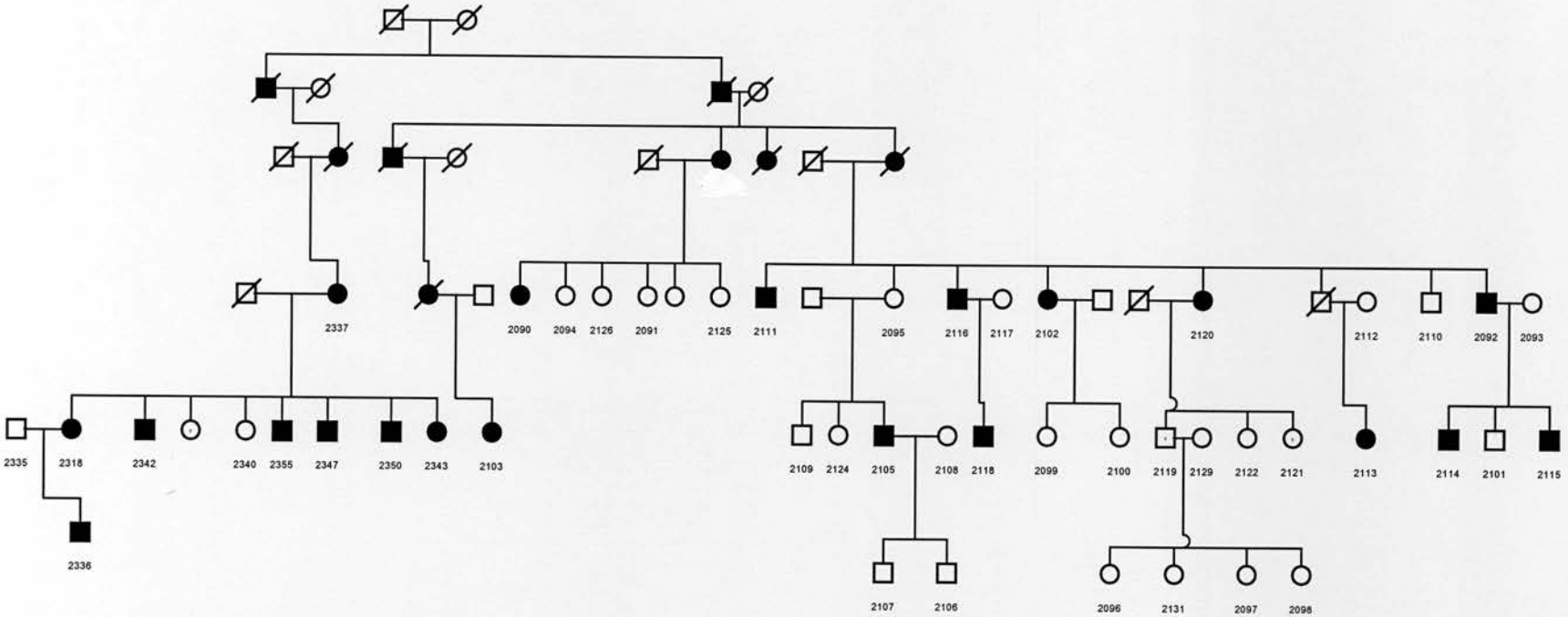


Figure 3.1 Family tree of family 9,
The number under the symbol is the reference number of each individual, used in table 3.1.
The black symbols denote affected status on clinical history and examination, as discussed in the text.

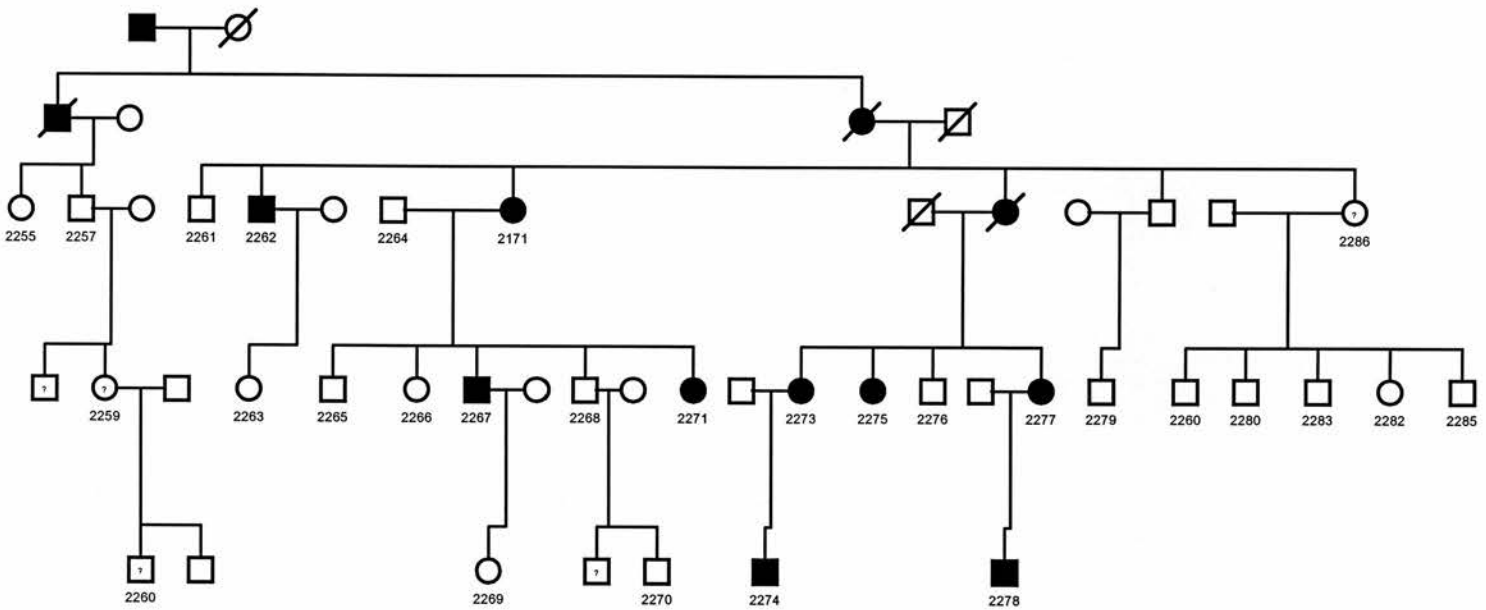


Figure 3.1 continued. Family tree of family 8.

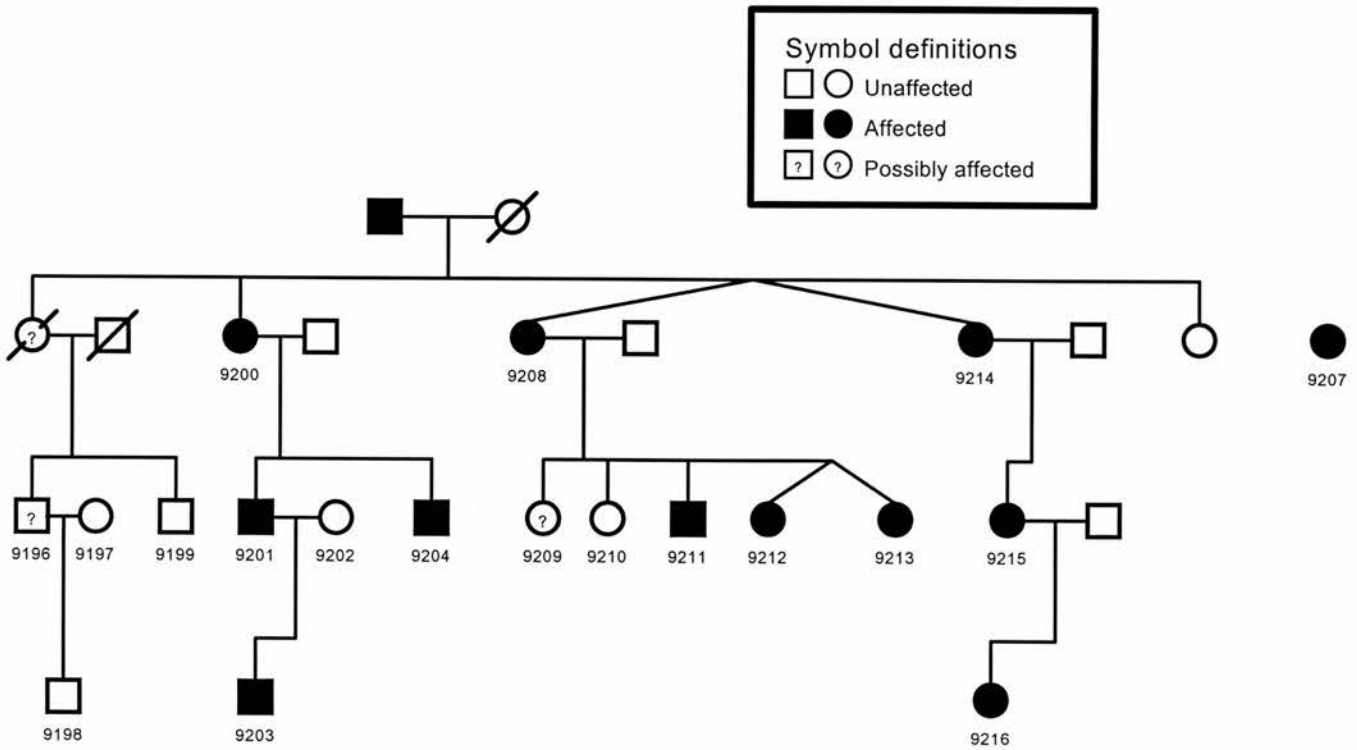
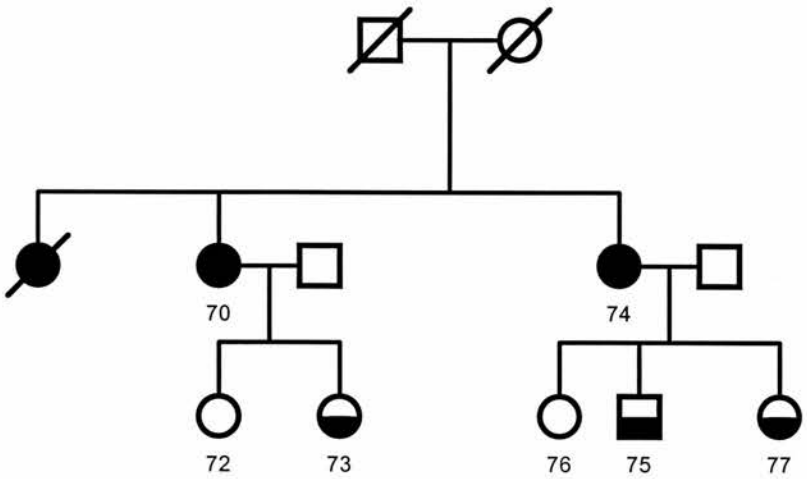


Figure 3.1 continued. Family tree of family 5.
Individual 9207 was thought to be a distant relative, currently living in the same region as other family members.

Family 6



Family 7

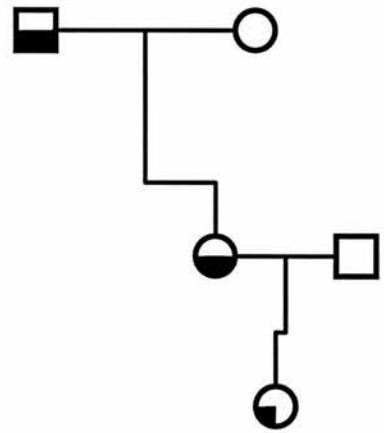
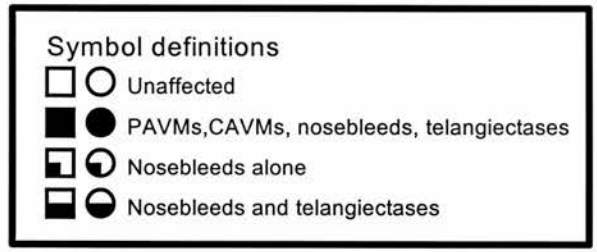


Figure 3.1 continued. Family tree of families 6 and 7.



Family Number	Number Affected	Epistaxis (Monthly)	Telangiectasis (2 lesions)	G-I haemorrhage (Suspected)	Hepatic	CAVM	PAVM
5	**12**	12	9	3	1	0	0
6	**5**	3	5	0	0	1	2
8	**9**	9	7	1	1	0	2
9	**20**	18	16	3	0	0	0
Total	**46**	42	37	7	2	1	4

Table 3.1 showing number of individuals affected within each family, and number of affected individuals with each disease feature. The diagnostic criteria used to define affected status were the presence of two of

(a) affected first degree relative.
(b) Monthly Epistaxis
(c) Characteristic Telangiectases
(d) Presence of a PAVM.

B. Markers for the endoglin region on chromosome 9

Family 8 was shown to be linked to chromosome 9 by other workers. A mutation was subsequently identified in the endoglin gene (a 4 base pair deletion in the extracellular region ΔCAGA 1078-81).

Markers D9S61 and D9S195 flank the endoglin region on chromosome 9 and are situated 11cM apart. D9S63 lies within 1cM of endoglin, and close to D9S61. These markers were used to test for linkage to the endoglin region in families 5 and 6.

Family 6 showed no recombination with D9S63 and with a two point LOD score of 1.3 at θ=0. With 5 informative meioses in the analysis, the maximum possible LOD score expected would be 1.5. A LOD score of 1.3 therefore suggested linkage to 9q34.

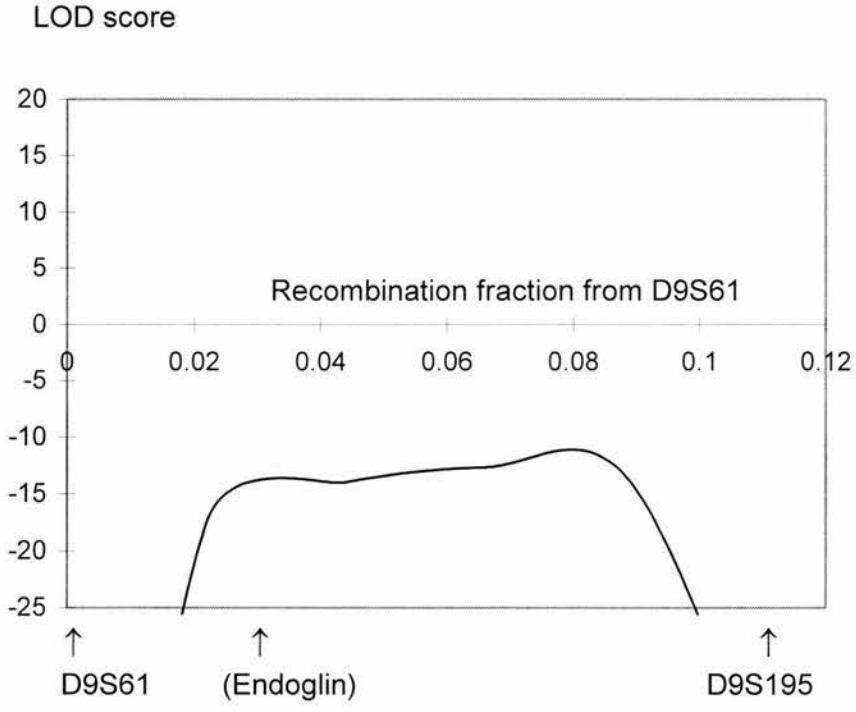
2 point linkage analysis for family 5 gave negative LOD scores with these markers. Linkage of family 5 to the region between D9S61 and D9S195 was formally excluded with a LOD score of less than -2 across the region on multipoint linkage analysis.

LOD scores for these analyses are given in **table 3.2a**

Locus	Family	Recombination Fraction (θ)					
		0.0	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**
D9S61	5	-4.7	-1.7	-0.39	0.05	0.29	0.25
D9S195	5	-14	-5.1	-3.0	-2.1	-1.1	-0.57
D9S63	6	1.3	1.3	1.3	1.3	1.2	1.0

Table 3.2a. (above) Two point LOD scores for families 5 and 6, for markers close to the endoglin locus. Generated using the MLINK program. Markers D9S61 and D9S195 are 11cM apart, spanning the endoglin locus. D9S63 is within 1cM distance of endoglin.

(below) Results of multipoint analysis for family 5. The LOD score is less than -2 across the entire interval, excluding the region between D9S195 and D9S61. The approximate position of Endoglin taken from GDB is marked.



C. Initial Linkage to the HHT2 Locus on Chromosome 12

A second locus in the peri-centromeric region of chromosome 12 was identified by Vincent (Vincent et.al 1995, with linkage to this region also seen in the previously published American families 17 and 33 (Johnson et.al. 1995).

Markers from this region were analysed in British families 1,2 and 3 previously described (Porteous et.al. 1994), and the newly ascertained families detailed in this chapter, families 5 and 9. LOD scores for these markers are given in **table 3.2b**, with map positions of the markers used.

Families 3 and 9 each individually gave two point LOD scores greater than 3 for markers D12S339 or D12S368, confirming linkage to the chromosome 12 locus.

Family 2 gave a LOD score of 2.0, suggesting linkage to this locus.

Initially family 5 failed to give either significantly positive or negative LOD scores with these markers on chromosome 12. The most likely cause for this was felt to be inaccurate diagnosis of individuals. Using more rigid diagnostic criteria, only classifying those with monthly nosebleeds and 2 or more telangiectases as affected, and including only affected individuals, the maximum LOD score of 1.9 was found with markers D12S339 at θ=0. The

Distance		8cM	3cM	3cM	
Marker		**345**	**85**	**339**	**368**

Family Number	θ	**D12S345**	**D12S85**	**D12S339**	**D12S368**
1	0.0	-2.8	-	-	-2.8
	0.01	0.09	-	-	-0.16
	0.05	0.67	-	-	0.47
	0.1	**0.81**	-	-	**0.65**
	0.2	0.76	-	-	0.63
	0.3	0.55	-	-	0.45
2	0.0	-6.8	-	-	**2.0**
	0.01	-1.5	-	-	2.0
	0.05	-0.29	-	-	1.8
	0.1	0.09	-	-	1.5
	0.2	**0.25**	-	-	1.1
	0.3	0.16	-	-	0.61
3	0.0	-10	**1.5**	**2.2**	**3.2**
	0.01	-3.6	1.5	2.2	3.1
	0.05	-1.9	1.3	2.0	2.8
	0.1	-1.14	1.14	1.7	2.4
	0.2	-0.43	0.78	1.2	1.6
	0.3	-0.15	0.43	0.70	0.84
9	0.0	-	-	-	-
	0.01	-2.1	0.71	3.8	-0.70
	0.05	0.27	0.93	**3.9**	-0.20
	0.1	1.2	**1.0**	3.8	0.07
	0.2	**1.6**	0.97	3.2	0.21
	0.3	1.3	0.75	2.2	0.17
	0.4	0.66	0.42	1.0	0.07
5	0.0	-	-7.3	-1.7	-3.1
	0.01	-	-2.8	-0.42	-1.7
	0.05	-	-0.88	0.74	-0.47
	0.1	-	-0.17	**1.1**	-0.02
	0.2	-	0.27	1.0	0.27
	0.3	-	**0.28**	0.71	**0.29**
5*	0.0	-	-2.7	**1.9**	**0.28**
	0.01	-	-0.16	1.8	0.27
	0.05	-	0.40	1.7	0.25
	0.1	-	**0.53**	1.5	0.22
	0.2	-	0.47	1.1	0.17
	0.3	-	0.26	0.61	0.12

Table 3.2b Pairwise LOD scores between HHT and markers D12S345, D12S85, D12S399 and D12S368. Approximate map distances for the markers based on Gyapay et.al. (1994) are given on the bar above the table. "-" marks a LOD score that has not been calculated. Zmax is marked in bold. The analyses were run using MLINK as described in materials and methods. Analysis with family 5 was run firstly using original diagnostic criteria, and then using more rigid criteria (marked as 5*) as described in the text.

total number of potentially informative meioses in this analysis was 7, leading to an expected maximum possible 2 point LOD score of approximately 2.1. It was therefore felt likely that family 5 was linked to HHT2 on chromosome 12.

D. Haplotype Analysis to Narrow the Candidate Interval for HHT2.

Figure 3.2 shows a simplified physical map of a 1.38Mb YAC contig spanning a part of the candidate region for HHT2. The locations of polymorphic markers, expressed sequence tags and known genes are given. Additional markers were analysed from this region to identify critical recombinations in families 2, 3 and 9, narrowing the candidate interval for HHT2. The haplotype analysis of these markers for families 3 and 9 is shown in **figure 3.3.**

Looking only at critical recombinations in affected individuals, family 2 showed a recombination between D12S347 and D12S297 placing HHT telomeric of D12S347. In family 67 a recombination between D12S359 and D12S368 placed HHT2 centromeric of D12S359. This reduced the candidate region to 4 centiMorgans. An unaffected individual in family 67 has a recombination between D12S368 and D12S297 placed HHT2 centromeric of D12S297 reducing the candidate interval further to

approximately 1 centiMorgan. These recombinations are shown diagramatically in **figure 3.4.**

The genes known to map within this region were the ALK-1 gene, an endothelially expressed serine-threonine kinase, and ALK-4, the ubiquitously expressed activin 1B receptor.

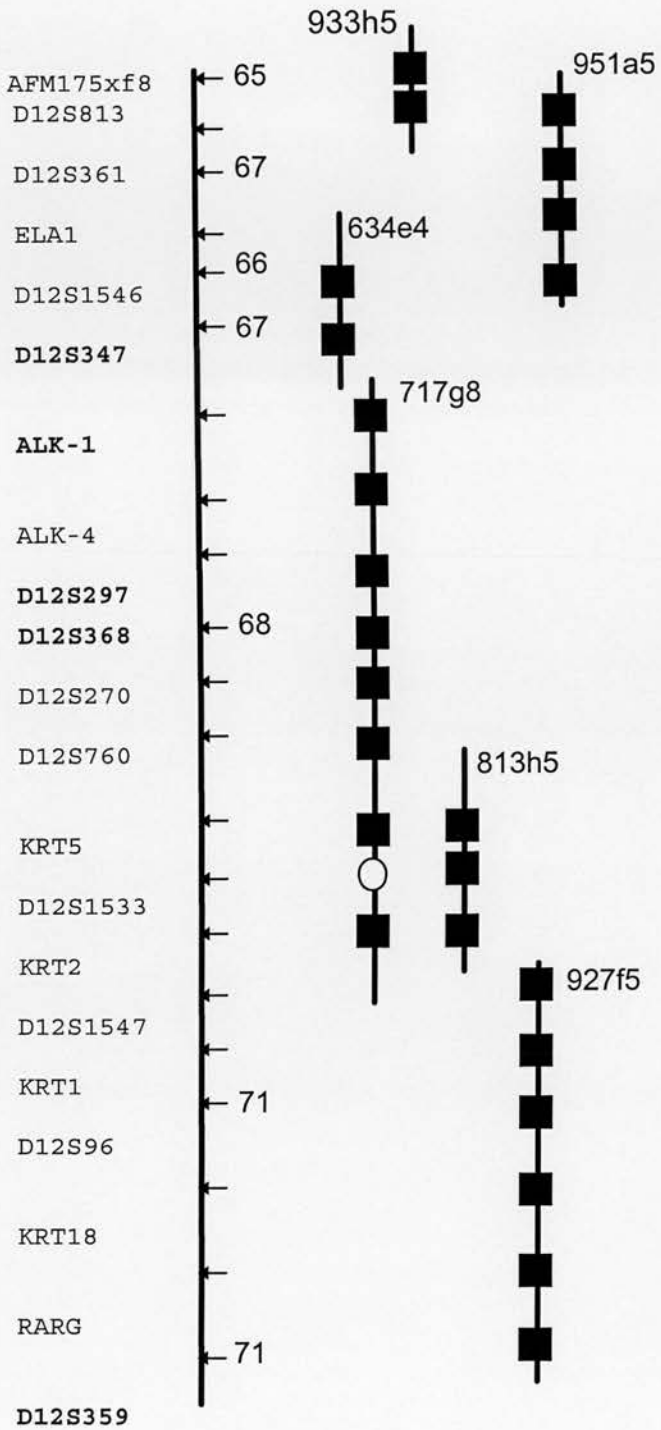


Figure 3.2 Physical map of a part of the HHT2 candidate region surrounding marker D12S368.(Supplied by Kucherlapati et.al.) The thin black line represents a portion of the long arm of chromosome 12 with the centromere at the top. The locus names are given on the left hand side. Markers used in this study are marked in bold. Arrows mark the position of each locus. Where a genetic distance for a marker is given, adjacent to an arrow, it is derived from the Genethon linkage map. the lines on the right hand side represent a YAC tiling path spanning the region. The number of the YAC is given adjacent to the line representing it. The black squares mark the position of each locus on the YACs. The white circle represents a marker not found, presumeably due to internal deletion of YAC 717g8.

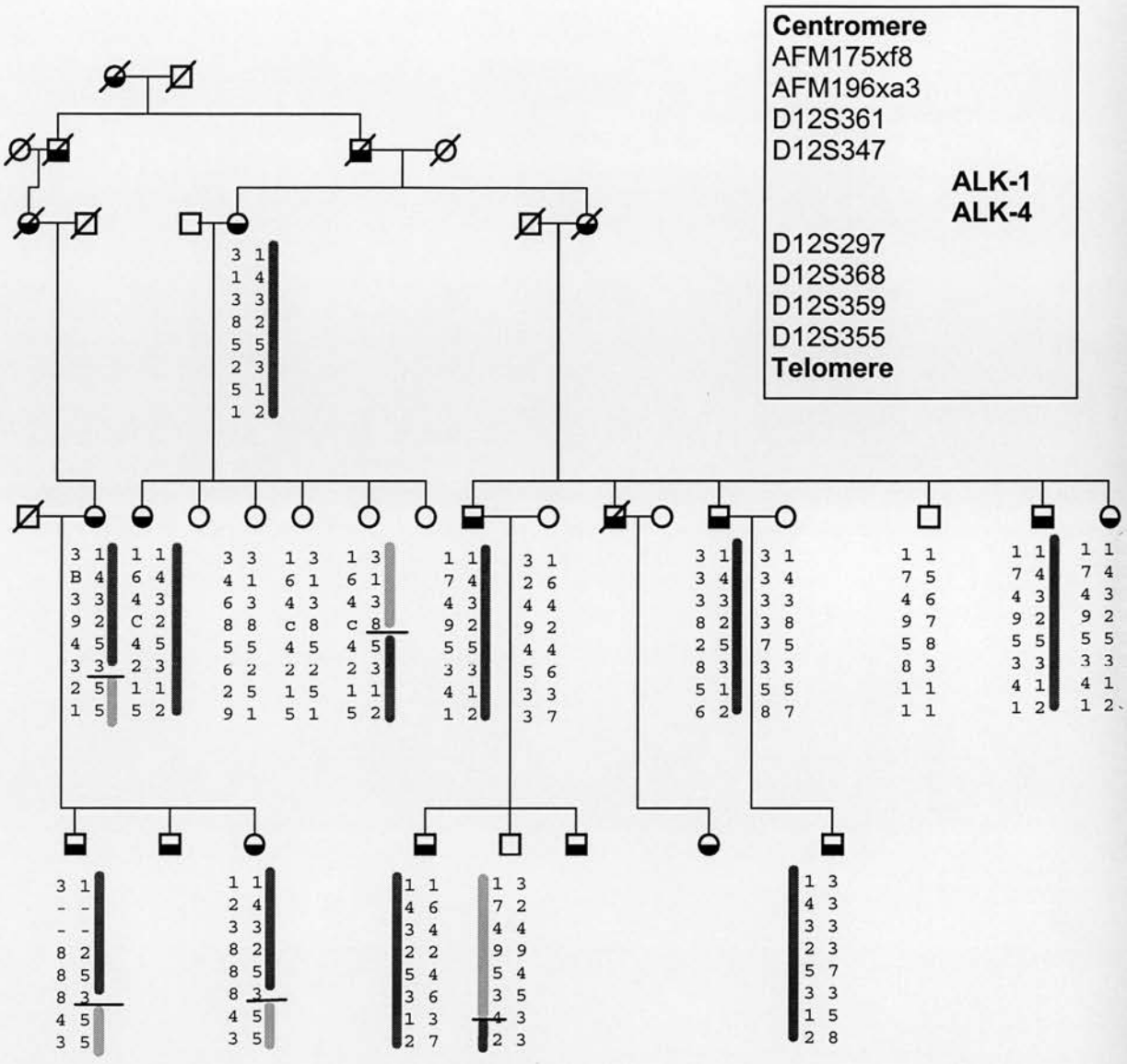


Figure 3.3a. Haplotype analysis of polymorphic markers from the Critical region of chromosome 12 in family 9. Only a proportion of individuals and haplotypes are shown for clarity. The letters or numbers represent alleles at each locus. A "-" is used where marker data was not available. The order of loci along chromosome 12 is given above.

The dark grey bars mark the haplotype linked to HHT2 in the family. Where a critical crossover ocurrs, the portion of the haplotype derived from the unaffected chromosome in the affected parent is marked as a light grey bar. Black lines mark the estimated position of the recombination at meiosis.

Critical crossover in affected individual iv.2 between D12s368 and D12S359 defines a 4cM critical interval for HHT2. A crossover in unnaffected individual iv.7 places HHT centromeric of D12S368, narrowing the candidate region to approximately 1cM.

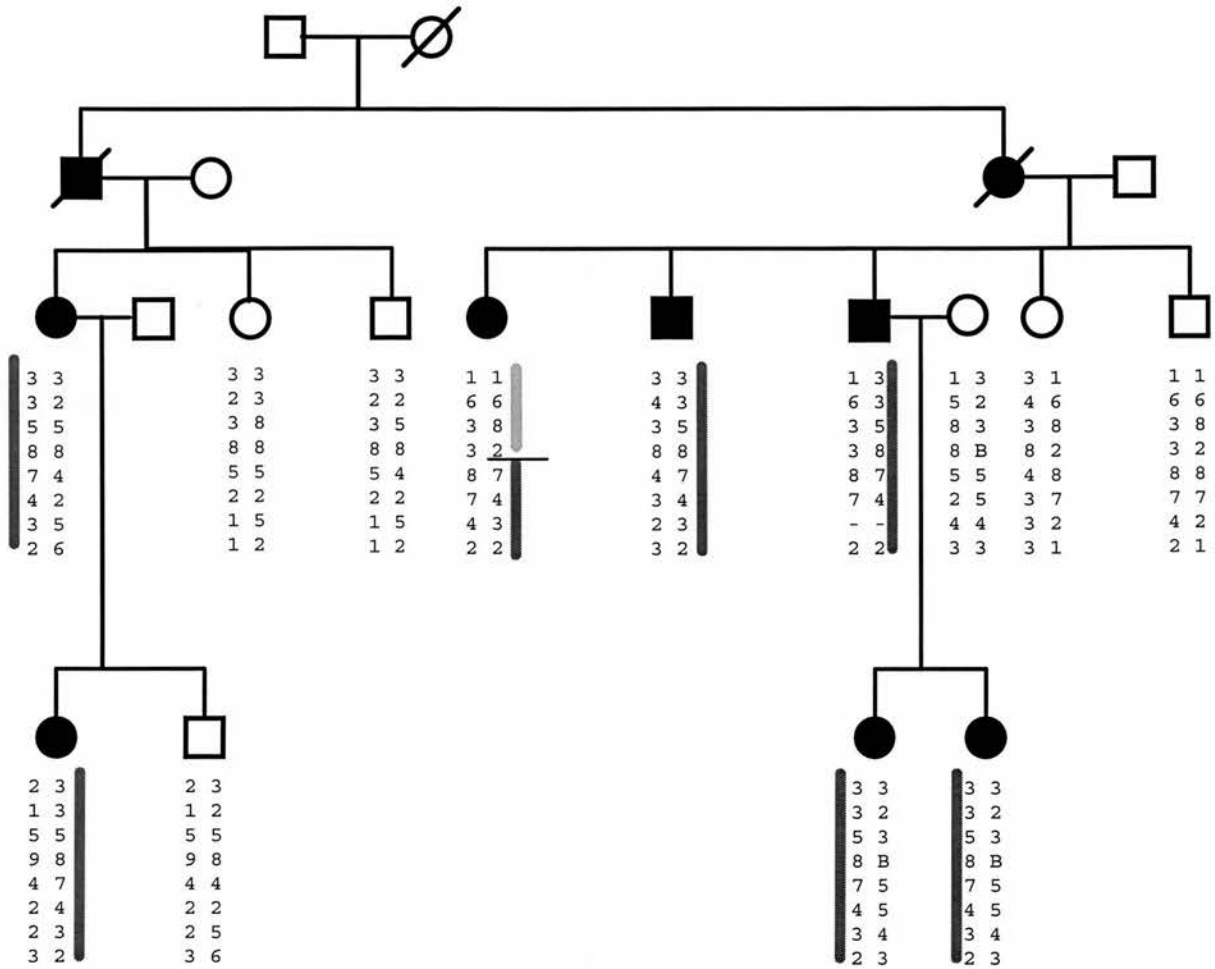


Figure 3.3b Simplified pedigree of family 3, showing haplotype analysis for the same markers as family 9. The critical recombination event in individual iii.5 between D12S347 and D12S 297 places HHT2 telomeric of D12S347.

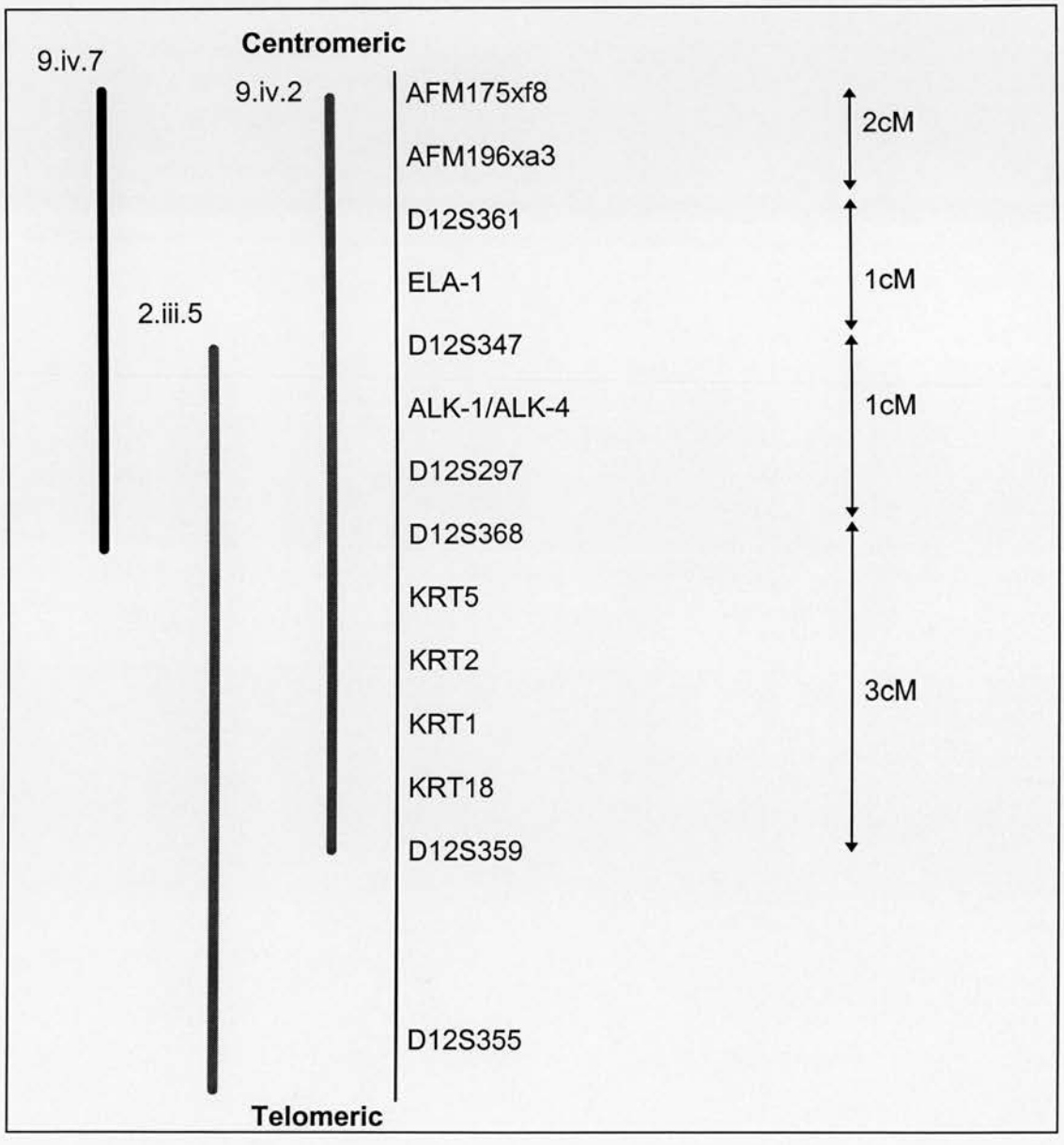


Figure 3.4. Diagram showing the critical region defined by the recombinations in family 9 individuals iv.2 and iv.7 and family 2 individual iii.11.

The thin black line represents the chromosomal segment with loci marked on the right hand side. The map distances given, marked in centiMorgans by the arrows are derived from the physical map of the region.

The broad lines indicate the chromosomal regions defined by the crossovers.

4. Investigation of genes mapping within the HHT2 Candidate Region.

Several genes mapped within the initially defined candidate interval for HHT2. As well as the known genes ALK-1, ALK-4 (the activin IB receptor) , and several keratin genes, there were two uncharacterised ESTs, WI-8355 and WI-6754. Expression of these in endothelial cells had previously been demonstrated by rtPCR. (D. Johnson unpublished data). Initially the Integrin-α5 gene was also felt to be a strong candidate for the HHT2 gene, as it mapped close to the region and is thought to play an important role in cell-cell and cell-matrix interaction of endothelial cells.

Investigation of ESTs WI-8355 and WI-6754

A. Using PCR to amplify directly from cDNA libraries.

The first approach attempted was to amplify fragments of the desired ESTs directly from an endothelial specific cDNA library. The cDNA libraries available were provided by Dr. D. Ginsburg (Ginsburg et.al. 1985). cDNA had been cloned bidirectionally into the EcoR1 site of λgt11. Both a library made from oligo-dT primed cDNA (designated EC-1) and one from random primed cDNA (designated EC-2) were used.

Repeated attempts to perform a hemi-nested PCR from EST specific primers to a vector specific primer were made using different dilutions of

starting template library. Between 10^6 and 10^{10} clones were used as template.

No product fragments for WI-8355 were generated repeatably or at sufficient concentration to justify sequencing.

Using primers made to EST WI-6754, several fragments were generated which were sequenced. One fragment had a significant open reading frame of 96 codons. Blast search of this open reading frame at the nucleotide and protein level revealed that it was the human c-yes-1 gene known to map to chromosome 18. Further inspection of the sequence identified an EcoR1 site between the open reading frame and the known WI6574 sequence. This suggested that the clone which amplified was chimaeric and contained 2 inserts, one of which was c-yes-1. No other fragments were generated reliably or had a significant open reading frame.

Owing to these problems, this approach was abandoned in favour of directly screening the cDNA libraries.

B Screening the cDNA libraries for WI-6754 and WI-8355

After titreing, the EC-1 and EC-2 libraries were plated out at a density of $2X10^4$ per 15cm circular plate. 10 plates were prepared for each. If the

proportion of the desired RNA species was as low as 1 in 50,000, screening these 10 filters would be expected to yield 4 clones.

Screening with WI-6754 identified no positive clones. Screening with WI-8355 identified multiple clones, 20 of which were isolated with two further rounds of plaque purification.

Clones isolated were then checked by PCR between EST specific primers and vector specific primers. As the library was cloned non-directionally, both orientations had to be tried. Three clones which gave consistent results in this assay were chosen for further sequencing. The PCR results for these clones, 8355-2, 8355-11 and 8355-12 are shown in **figure 4.1.**

Sequencing 150 bp of the 3' end of these fragments yielded identical sequence to the published EST. 120 to 220 base pairs of sequence was generated from the 5' end of each of the 3 clones. None of these short sequences overlapped. However, each contained an open reading frame in the expected orientation.

The nucleotide sequence and deduced protein sequence was used to search Genbank using the Blastn and Blastp programmes.

At the nucleotide level, clone 12 was identical to a 500 base pair random human cDNA clone ym50co2.r1 (gb|H24106). At the protein level, there

was extensive homology of predicted amino acid sequence from the clone 12 fragment to helicase proteins from many different species. However, no exact matches with human proteins were identified. Sample alignments are shown in **figure 4.2.** Re-running the database search with the entire sequence of ym50co2.r1 (Last performed April 1997) did not identify any complete mRNA sequences.

The 5' end of clones 2 and 11 had weak homology to human U1 small nuclear ribonuclear protein at the protein level.

Helicases have been identified as mutated in the alpha-thalassaemia and mental retardation syndrome (Ion et.al. 1996). They are also implicated in disorders of DNA repair and cancer predisposition such as Bloom syndrome, Werner Syndrome and Xeroderma pigmentosum (Epstein and Motulsky 1996). This did not seem a likely class of gene for a candidate for HHT2. At this point it became known that WI-8355 also fell outwith the most recently defined candidate interval for HHT2, and characterisation was therefore suspended.

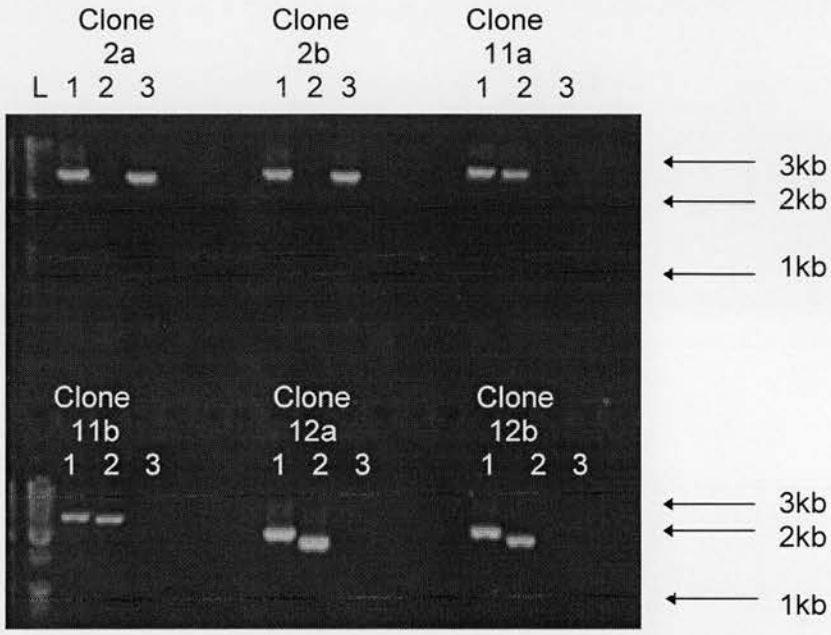


Fig 4.1. PCR results from clones 2, 11 and 12. 2 clones were analysed for each. In all cases, the reactions were loaded

Lane 1. PCR using vector specific primers λgt11 forward and λgt11 reverse.
Lane 2. PCR using λgt11 forward primer and an EST specific reverse primer.
Lane 3. PCR using λgt11 reverse primer and an EST specific reverse primer.

For each clone there is a marginally larger product in lane 1 than in 2 or 3, as would be expected. Each clone also gives a product in lane 2 or 3 but not both, as predicted from experimental design.

Sizes are derived from Gibco-BRL 1kb ladder in Lane L. (indistinct on this gel).

```
Clone 12.............RNRDIIGVAETGSGKTAAFLIPLLVWITTLPKIDRIEESDQGPYAIILA
C.elegans f01f7.1Q....NRDVIGVAETGSGKTAAFLLPLLVWITSLPKMERQEHRDLGPYAIIMA
Slime-Mould Helicase...RDILGIAETGSGKTCAFVIPMLIYISKQPRLTKDTEAD GPYALVMA
Human p68 RNA Helicase..DMVGVAQTGSGKTLSYLLPAIVHINHQPFLER
```

Figure 4.2

Alignment of 49 amino acids from the translated sequence from the 5' end of clone 12. This stretch had significant homology ($p<10^{-4}$) to 22 published helicases. However, at the time of doing the blast search no full length human cDNA had been identified.

5. The Structure of the Activin Receptor Like Kinase 1 Gene.

A. A human genomic PAC clone containing the ALK-1 gene.

From the YAC contig described in chapter 4, it was anticipated that the ALK-1 gene was contained within CEPH YAC 717_g_8. Southern blots of digested 717_g_8 YAC DNA probed with an ALK-1 cDNA probe showed no specific hybridisation. A specific PCR reaction designed to amplify a 120bp fragment from the 5' UTR of the ALK-1 gene also failed to give a product from the 717_g_8 YAC DNA template. (Data not shown).

The human genomic PAC library supplied by Pieter de Jong (Ioannou et.al. 1994) was screened. Using the same PCR described above which specifically amplifies a 120bp fragment from the 5 UTR of the human ALK-1 gene, positive results were found for DNA pools E3, E5 and E9 corresponding to co-ordinates 5X3, 5Y1 and 5Z1 of the 3 dimensional grid into which the plates were arrayed. This identified a single plate number 265 as containing a genomic clone corresponding to ALK-1.

An impression of plate 265 was stamped into LB agar with kanamycin, followed by overnight incubation at 37 ^{O}C. Clone DNA was transferred to Hybond N membrane as previously described. Hybridisation to a radiolabelled ALK-1 cDNA probe identified a single clone as containing ALK-1 sequence. This clone at position c-14 was designated **GA-1.**

B. GA-1 has an insert of 120kb and contains two Not 1 sites

DNA from the clone GA-1 was digested with a variety of restriction endonucleases and resolved using both conventional electrophoresis and pulsed field gel electrophoresis.

Not 1 digestion of the clone yielded fragments of approximately 60 kb, 45kb and 15kb as well as a vector fragment of 20kb when resolved on pulsed field gel electrophoresis, as shown in **figure 5.1**. This gave an estimate of the total insert size as around 120kb.

A standard southern blot of Bgl II, EcoR1 and HindIII digests of GA-1 was probed with the total ALK-1 cDNA coding region. The results are shown in **figure 5.2** and summarised in **table 5.1**. The fragment sizes correspond with those found on a human genomic southern blot. These results suggest that the coding region of the ALK-1 gene contained within GA-1 falls within less than 20kb of genomic DNA.

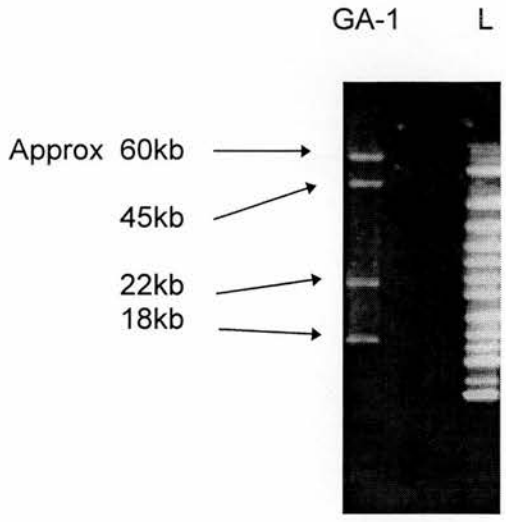


Figure 5.1 Pulsed field gel electrophoresis of a NOT1 digest of PAC clone GA-1. The ladder (Gibco high molecular weight marker) is on the right hand side. Fragment sizes of the digest are given on the left hand side.

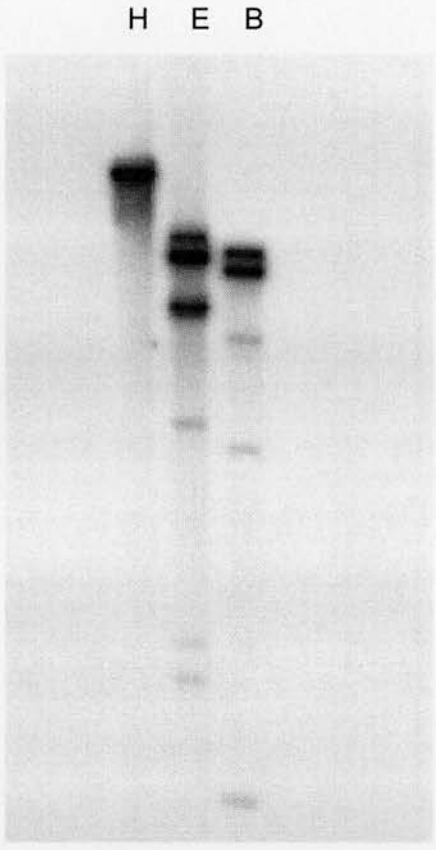


Figure 5.2 Southern blot of digests of the PAC clone GA-1 probed with the complete cDNA for ALK-1. Sizes of the fragments for the HinDIII (H), EcoRI (E) and Bgl II(B) digests are given in **table 5.1** below

Enzyme Digest	Size of Fragments (kb)	Total Size (kb)
Bgl II	7, 5.5, 3.8, 2.3, 0.8	19.4
EcoR I	8, 6, 4.5, 2.5, 1.5, 1.7	24.2
HinD III	(22)	22

Table 5.1. Approximate fragment sizes (in kilobases) indicated by the bands identified in the autoradiogram shown in **figure 5.2**. The size of the HinD III fragment is given in parentheses, as it was calculated by pulsed field gel electrophoresis rather than from the autoradiogram shown.

C. Subcloning of GA-1

A Bgl II digest of GA-1 DNA was subcloned into the BamH1 site in pBluescript (Stratagene). 24 clones that hybridised to ALK-1 cDNA were selected. Double digestion with EcoR1 and Not1 identified 3 different groups of clones. The fragment sizes generated for each of these clones by this digestion are given below.

Group 1 : Insert digested to give fragments of 3kb and 2kb.
Group 2 : Insert digested to give fragments of 4kb and 1.4kb.
Group 3 : Insert digested to give fragments of 4.2kb, 2.2kb and 1.8kb.

A southern blot of these clone digests was sequentially probed with specific primer 460, located at the 5' end of the coding region of ALK-1 and primer 505 from the coding 3' end the gene. Clones from group 1 hybridised specifically to primer 460, and those from group 2 hybridised specifically to primer 505. Clones from group 3 were shown to hybridise specifically to a probe from the 5' untranslated region of ALK-1.

Subclones 2b (group 1), 1a (group 2) and 6d (group 3) were selected as representative of their groups. The position of each clone in relation to the ALK-1 gene is shown as a part of **figure 5.3.**

In addition to digest patterns, the clones 6d and 2b were characterised further by PCR from an ALK-1 specific primer to the stratagene T7 and T3 primers on either side of the multiple cloning site of the pBluescript vector used. This demonstrated that clone 6d contained 2.2kb of sequence between the 3' end of ALK-1 exon 1 and the T7 primer. Clone 2b contains approximately 3kb of sequence between the 5' end of ALK-1 exon 2 and the T7 primer. (data not shown).

D. Determination of Intron Position in the ALK-1 Gene.

Intron position in the ALK-1 gene was determined using two techniques, exon to exon PCR and sequencing from genomic subclones 2b, 1a and 6d.

Using PCR, it was possible to amplify the genomic DNA between one exon and the next by using the DNA from the GA-1 clone as template. Primers designed to the published cDNA sequence of ALK-1 were used. An example is given in **figure 5.4**. Wherever a fragment generated was longer than that expected from the known cDNA sequence, an intron was identified. This was confirmed by sequencing, the exact intron-exon border being identified by a consensus splice sequence followed by divergence from the established cDNA sequence. **Table 5.2** gives the fragment sizes generated from genomic PCR compared to the sizes expected from PCR of the cDNA sequence. This data was used to construct **figure 5.3**

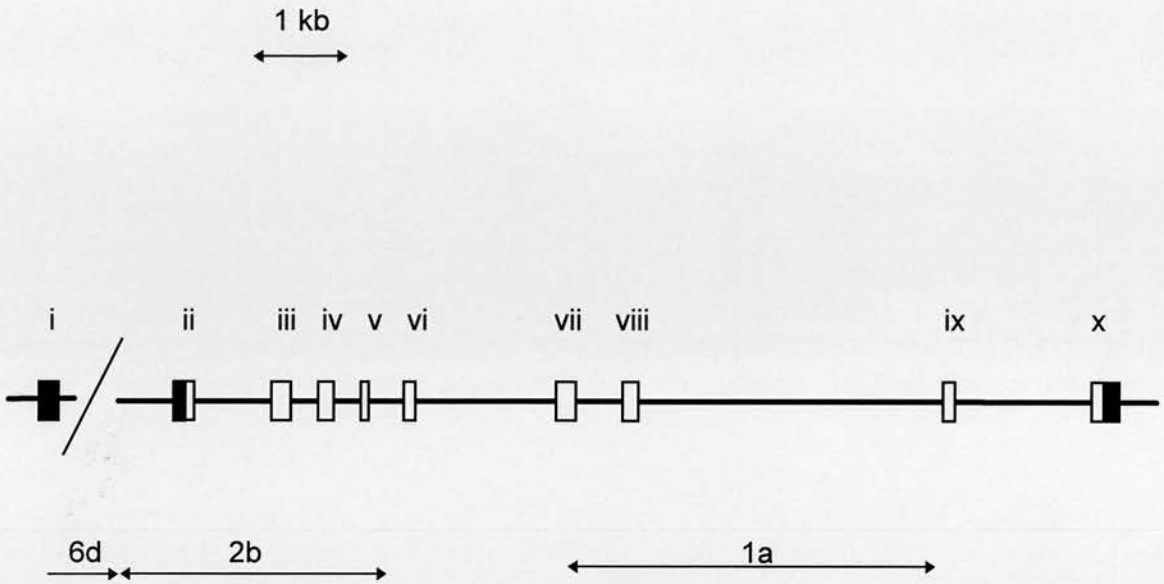


Figure 5.3. The Genomic structure of ALK-1. Drawn approximately to scale.

The exons are numbered with Roman numerals

Exon, untranslated region

Exon, translated region

Intron

Region spanned by genomic subclone

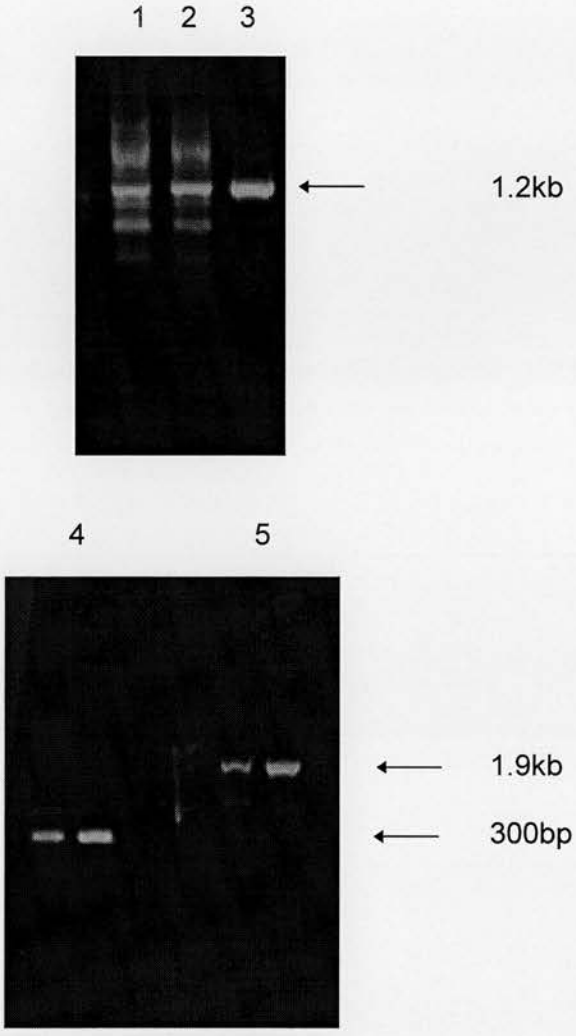


Figure 5.4 Examples of PCR between exons of ALK-1 used to identify intron position.

lanes 1 and 2 show product from PCR reactions using human genomic DNA as template, lane 3 used DNA from PAC clone GA-1 as template.

Lanes 1,2 and 3 show the fragment size of 1.2 kb generated using primers 460 and 283,this product would be expected to contain the second intron of ALK-1 (900bp). The product generated using the PAC clone as a template shows no additional bands and is easier to identify and sequence when compared to the product generated using genomic DNA as template.

Lane 4 shows the product generated using primers 148 and 280. This product contains the 3rd intron of ALK-1 which is 140bp in size. The product in lane 5 between primers 561 and 280 contains introns 2 and 3, 900bp and 140bp respectively.

Forward Primer	Position (exon No)	Reverse Primer	Position (exon No)	PCR Fragment Generated	Intron No (Size)
561	-41 (2)	460	171 (3)	1.2kb	2 (900bp)
561	-41 (2)	280	564 (5)	1.9 kb	2 (900bp) 3 (140bp) 4 (160bp)
124	473 (4)	280	564 (5)	250bp	4 (160bp)
124	473 (4)	490	724 (6)	700bp	4 (160bp) 5 (200bp)
489	659 (6)	125	1032 (7)	1.2kb	6 (800bp)
501	895 (7)	504	1179 (8)	750bp	7 (450bp)
505	1182(8)	506	1290(9)	2.9kb	8 (2.8)kb
507	1316(9)	560	1544(10)	1.7kb	9(1.5kb)

Table 5.2 Table giving primer pairs used in the PCR reactions to estimate intron sizes and positions. The primer identification number corresponds with that given in **table 2.1.**

The position of the primer in the cDNA sequence is given as the position of the first base of the primer in relation to the start codon of the published ALK-1 cDNA sequence. The approximate size of fragment generated using the primer pair on the GA-1 DNA template is given. This allowed the estimate of intron position and size given in the table.

The fragments generated were sequenced from a variety of primers to identify the intron-exon boundaries given in **Table 5.3**,and shown diagrammatically in **Figure 5.4**, as well as to generate the sequence data for the GENBANK submission given in **appendix A.**

Figure 5.5 gives the coding ALK-1 cDNA sequence with the position of each intron identified. For some intron-exon borders the sequence gained using this approach was not sufficiently clear. To get better sequence for these areas, further sequencing was performed directly using clones 1b and 2a as template.

Table 5.3 gives the sequence of each intron/exon boundary for the coding exons. The Genbank submission containing this sequence and extended sequencing into the introns is given in **appendix A**.

The 5' side of intron 6 does not have the usual consensus splice sequence. There is the splice donor sequence TAG/gcaag instead of the usual AG/gt.

From this it can be seen that the coding portion of the ALK-1 gene is contained within 9 exons and lies within just under 15kb of genomic DNA.

E. Further Characterisation of the Region Containing Exons 1 and 2

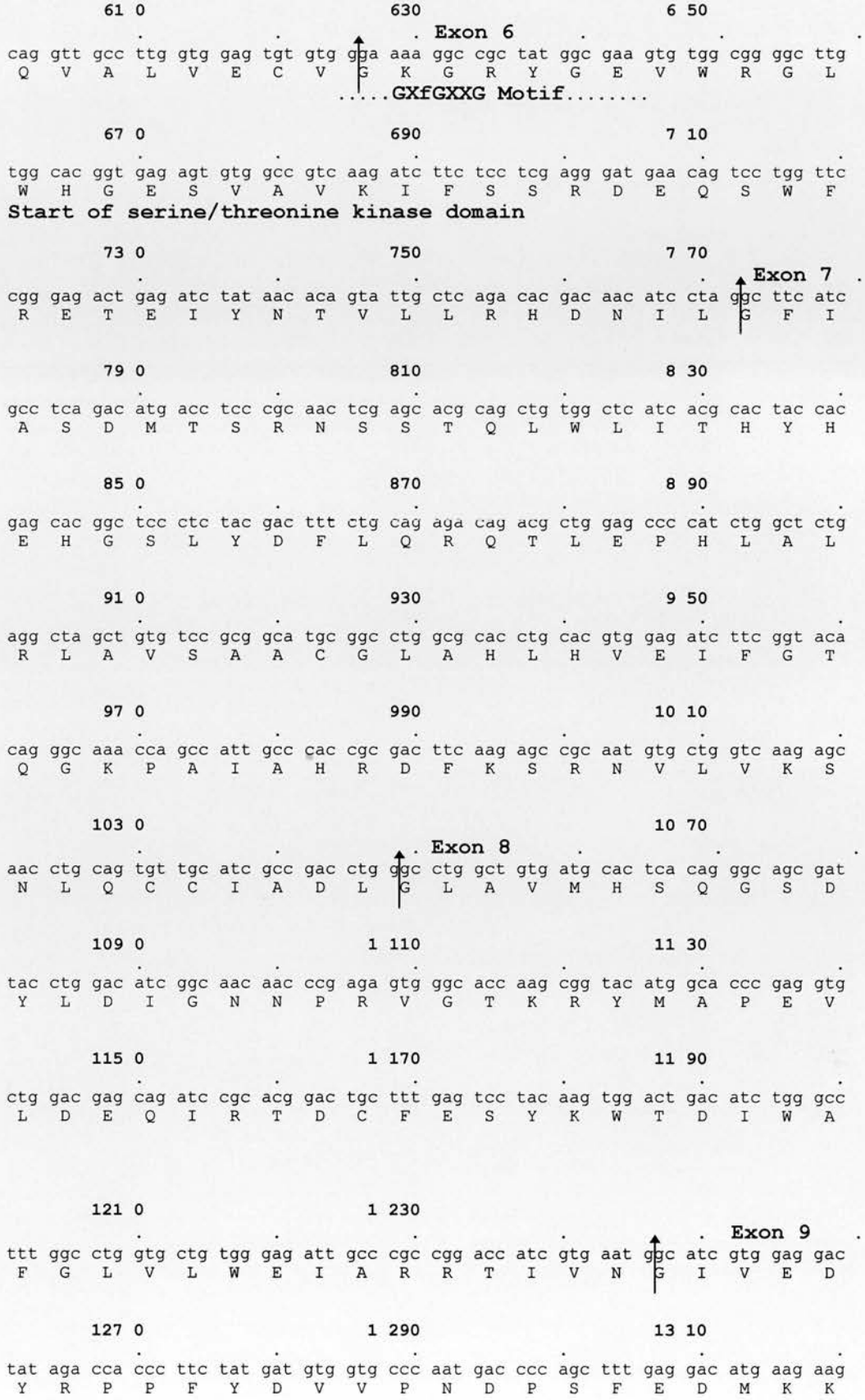
The 5' end of the ALK-1 gene that contains exon 1 and exon 2 is of particular interest as it would be expected to contain the promoter elements controlling the endothelial specific expression of ALK-1.

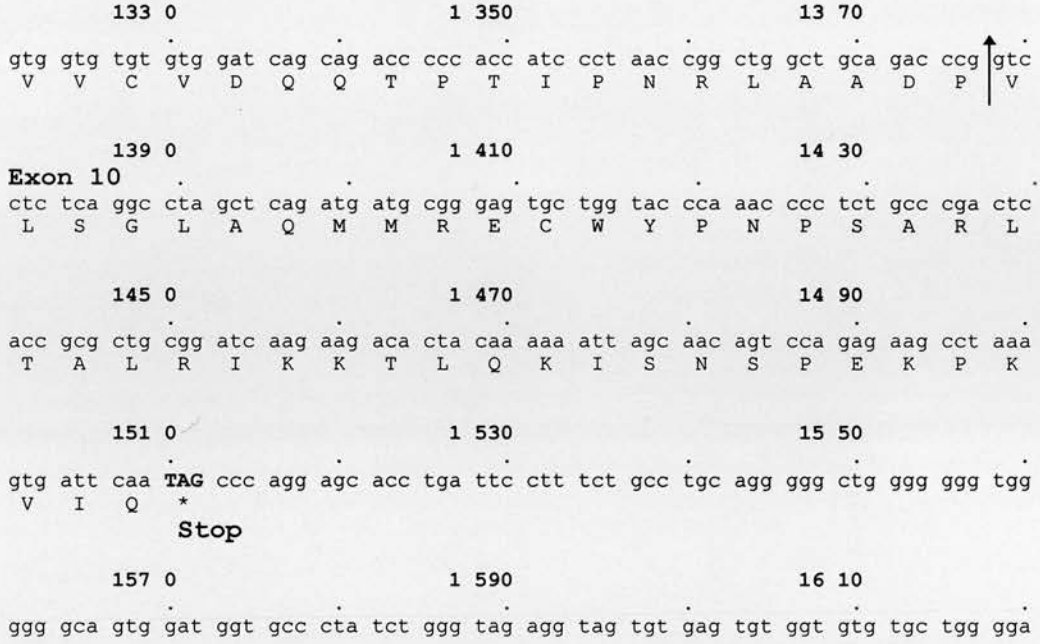
Figure 5.5

Annotated Sequence of the Coding Region of ALK-1. Exact Intron Position is marked by an arrow.

The Exon number is marked above the sequence. The start codon, and stop codon are marked in bold. The putative signal peptide for membrane localisation, putative transmembrane region, the GXGXXG motif, GS domain and start of the kinase domain are marked.

```
          1 0                        30                        50
Exon 2      .                         .                         .                         .
ATG acc ttg ggc tcc ccc agg aaa ggc ctt ctg atg ctg ctg atg gcc ttg gtg acc cag
 M   T   L   G   S   P   R   K   G   L   L   M   L   L   M   A   L   V   T   Q
Start                                      .....Putative Signal peptide.....

          7 0                        90                        1 10
  Exon 3    .                         .                         .                         .
gga gac cct gtg aag ccg tct cgg ggc ccg ctg gtg acc tgc acg tgt gag agc cca cat
 G   D   P   V   K   P   S   R   G   P   L   V   T   C   T   C   E   S   P   H

          13 0                       150                       1 70
            .                         .                         .                         .
tgc aag ggg cct acc tgc cgg ggg gcc tgg tgc aca gta gtg ctg gtg cgg gag gag ggg
 C   K   G   P   T   C   R   G   A   W   C   T   V   V   L   V   R   E   E   G

          19 0                       210                       2 30
            .                         .                         .                         .
agg cac ccc cag gaa cat cgg ggc tgc ggg aac ttg cac agg gag ctc tgc agg ggg cgc
 R   H   P   Q   E   H   R   G   C   G   N   L   H   R   E   L   C   R   G   R

          25 0                       270                       2 90
            .                         .                         .                         .
ccc acc gag ttc gtc aac cac tac tgc tgc gac agc cac ctc tgc aac cac aac gtg tcc
 P   T   E   F   V   N   H   Y   C   C   D   S   H   L   C   N   H   N   V   S

          31 0                       330                       3 50
            .    Exon 4 .             .                         .                         .
ctg gtg ctg gag gcc acc caa cct cct tcg gag cag ccg gga aca gat ggc cag ctg gcc
 L   V   L   E   A   T   Q   P   P   S   E   Q   P   G   T   D   G   Q   L   A

          37 0                       390                       4 10
            .                         .                         .                         .
ctg atc ctg ggc ccc gtg ctg gcc ttg ctg gcc ctg gtg gcc ctg ggt gtc ctg ggc ctg
 L   I   L   G   P   V   L   A   L   L   A   L   V   A   L   G   V   L   G   L
 ..................... Putative Transmembrane Domain .........................

          43 0                       450                       4 70
            .                         .                         .                         .
tgg cat gtc cga cgg agg cag gag aag cag cgt ggc ctg cac agc gag ctg gga gag tcc
 W   H   V   R   R   R   Q   E   K   Q   R   G   L   H   S   E   L   G   E   S

          49 0                       510                       5 30
            .                         .                         .          .Exon 5        .
agt ctc atc ctg aaa gca tct gag cag ggc gac acg atg ttg ggg gac ctc ctg gac agt
 S   L   I   L   K   A   S   E   Q   G   D   T   M   L   G   D   L   L   D   S

          55 0                       570                       5 90
            .                         .                         .                         .
gac tgc acc aca ggg agt ggc tca ggg ctc ccc ttc ctg gtg cag agg aca gtg gca cgg
 D   C   T   T   G   S   G   S   G   L   P   F   L   V   Q   R   T   V   A   R
                .....GS Domain.....
```


61 0
630
6 50
Exon 6
cag gtt gcc ttg gtg gag tgt gtg gga aaa ggc cgc tat ggc gaa gtg tgg cgg ggc ttg
Q V A L V E C V G K G R Y G E V W R G L
.....GXfGXXG Motif.....
67 0
690
7 10
tgg cac ggt gag agt gtg gcc gtc aag atc ttc tcc tcg agg gat gaa cag tcc tgg ttc
W H G E S V A V K I F S S R D E Q S W F
Start of serine/threonine kinase domain
73 0
750
7 70
Exon 7
cgg gag act gag atc tat aac aca gta ttg ctc aga cac gac aac atc cta ggc ttc atc
R E T E I Y N T V L L R H D N I L G F I
79 0
810
8 30
gcc tca gac atg acc tcc cgc aac tcg agc acg cag ctg tgg ctc atc acg cac tac cac
A S D M T S R N S S T Q L W L I T H Y H
85 0
870
8 90
gag cac ggc tcc ctc tac gac ttt ctg cag aga cag acg ctg gag ccc cat ctg gct ctg
E H G S L Y D F L Q R Q T L E P H L A L
91 0
930
9 50
agg cta gct gtg tcc gcg gca tgc ggc ctg gcg cac ctg cac gtg gag atc ttc ggt aca
R L A V S A A C G L A H L H V E I F G T
97 0
990
10 10
cag ggc aaa cca gcc att gcc cac cgc gac ttc aag agc cgc aat gtg ctg gtc aag agc
Q G K P A I A H R D F K S R N V L V K S
103 0
10 70
Exon 8
aac ctg cag tgt tgc atc gcc gac ctg ggc ctg gct gtg atg cac tca cag ggc agc gat
N L Q C C I A D L G L A V M H S Q G S D
109 0
1 110
11 30
tac ctg gac atc ggc aac aac ccg aga gtg ggc acc aag cgg tac atg gca ccc gag gtg
Y L D I G N N P R V G T K R Y M A P E V
115 0
1 170
11 90
ctg gac gag cag atc cgc acg gac tgc ttt gag tcc tac aag tgg act gac atc tgg gcc
L D E Q I R T D C F E S Y K W T D I W A
121 0
1 230
Exon 9
ttt ggc ctg gtg ctg tgg gag att gcc cgc cgg acc atc gtg aat ggc atc gtg gag gac
F G L V L W E I A R R T I V N G I V E D
127 0
1 290
13 10
tat aga cca ccc ttc tat gat gtg gtg ccc aat gac ccc agc ttt gag gac atg aag aag
Y R P P F Y D V V P N D P S F E D M K K


133 0
1 350
13 70
gtg gtg tgt gtg gat cag cag acc ccc acc atc cct aac cgg ctg gct gca gac ccg gtc
V V C V D Q Q T P T I P N R L A A D P V
139 0
1 410
14 30
Exon 10
ctc tca ggc cta gct cag atg atg cgg gag tgc tgg tac cca aac ccc tct gcc cga ctc
L S G L A Q M M R E C W Y P N P S A R L
145 0
1 470
14 90
acc gcg ctg cgg atc aag aag aca cta caa aaa att agc aac agt cca gag aag cct aaa
T A L R I K K T L Q K I S N S P E K P K
151 0
1 530
15 50
gtg att caa TAG ccc agg agc acc tga ttc ctt tct gcc tgc agg ggg ctg ggg ggg tgg
V I Q *
Stop
157 0
1 590
16 10
ggg gca gtg gat ggt gcc cta tct ggg tag agg tag tgt gag tgt ggt gtg tgc tgg gga

Exon	3' Splice Junction	5' Splice Junction
2	*tgctcctctctgcag*GGACC**ATG**ACCTTGGG	GTGACCCAGG*gtgagtactgggggagcagtta*
3	*agcttccggtgtgtcttccag*GAGACCCTGT	GGTGCTGGAGG*gtacgtccagctgccctagca*
4	*cagtgtccccctccctcag*CCACCCAACCTC	CAGCATGTTGGGG*gtatgggcctggggacctg*
5	*cctctccgtacccccag*GACCTCCTGGACAG	GTGGAGTGTGTGG*gtgagcagtgggtgagccc*
6	*cttcccctctggccatcag*GAAAAGGCCGCT	ACGACAACATCC<u>TAG*gcaaggggagaggccag*</u>
7	*caacctttctgcacacag*GCTTCATCGCCTC	ATCGCCGACCTGG*gtgagccgggcggggcagg*
8	*caggcctcacccccacag*GCCTGGCTGTGAT	CCGGACCATCGTGAATG*gtgagggcccaccct*
9	*ccattctccatttccag*GCATCGTGGAGGAC	GGCTGGCTGCAGACCCG*gtgaggcctctgctg*
10	*ctctctcccaacccccag*GTCCTCTCAGGCC	GAAGCCTAAAGTGATTCAA**TAG**CCCAGGAGCA

Table 5.3 Showing the splice junctions at the 5' and 3' sides of each of the coding exons of ALK-1. Exonic sequence is in upper case, intronic sequence in lower case. The 5'splice site in exon 2 that leads to the cDNA sequence published by ten-Djike is given. The start and stop codons are marked in bold. The non-consensus splice junction on the 3' side of exon 6 is underlined.

The two cDNA sequences published for ALK-1 have different 5' untranslated regions. The sequence published by Atissano et.al. is contained entirely in exon 2 upstream from the expected start codon. The sequence published by ten Djike et.al. would arise as a result of exon 1 splicing into exon 2 at a consensus splice sequence 7 base pairs upstream from the start codon.

Specific primers, JB3 and JB4 designed to the cDNA sequence from exon 1 were used to amplify the intron between exon 1 and exon 2, giving a 5.5kb fragment **(shown in figure 5.6**). This fragment contains a single Bgl II site confirming that the Bgl II subclones described above, *6d* containing exon 1 and *2b* containing exon 2 are adjacent. It had previously been established by PCR that the insert of clone 6d contained approximately 2.2kb of sequence 3' of exon 1, and clone 2b contained approximately 3kb of sequence 5' of exon 2

Partial restriction mapping of this region was also carried out using pulsed field gel electrophoresis. Using the known Not 1 restriction site in exon 1 and a series of double digests combining Not 1 digestion with HinDIII, EcoR1 and BamH1 digests were resolved in the 5 to 50kb size range by pulsed field gel electrophoresis and Southern blotted, before being hybridised to specific oligonucleotides complementary to exon 1 and exon 2.

The oligonucleotides JB3 (exon 1, 3' of Not 1 site) and 647 (in exon 2) both hybridised specifically to the largest Not 1 fragment of approximately 50kb as well as a to the largest HinD III fragment of 22kb. The Southern blot of the Not1/HinDIII digest showed specific hybridisation to an 8kb fragment.

Digestion of clone 6d with Not 1, EcoR1 or both, gave the fragment sizes shown in **table 5.2.** Combined with the estimation of distance between exon 1 and the 3' end of the insert of 2.2kb derived by PCR and described above, this gave the restriction map shown in **figure 5.7.**

F. Conservation of Intron Position in the Kinase Domain of ALK-1.

The full genomic structure of one other type I serine-threonine kinase receptor has been published. This is the TSK-7L receptor isolated from mouse. Comparison of the splice junctions at the protein level between ALK-1 and TSK-7L are shown in **figure 5.8.**

All 3 splices between exons 5,6,7 and 8 all fall in exactly the same position, between 1st and second base of a glycine codon. The splice between exon 8 and 9 falls after the 1st base of a codon for Isoleucine, another neutral amino acid.

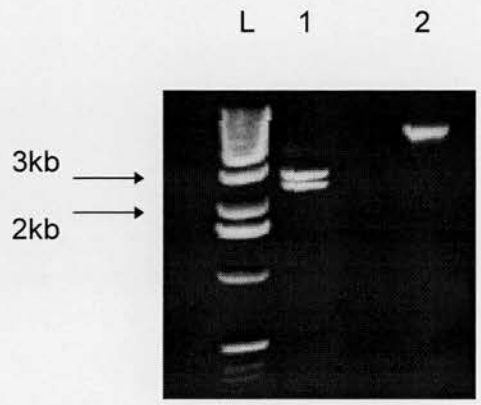


Figure 5.6 PCR product amplified using primers JB3 and JB4, amplified using clone GA-1 as template. Lane 2 shows the 5.5kb undigested product. Lane 1 shows the Bgl II digest which shows a single site in the fragment, lying between exon 1 and exon 2 as described in the text.

The ladder in lane L is Gibco BRL 1kb ladder. The size bands for 2kb and 3kb are marked with an arrow.

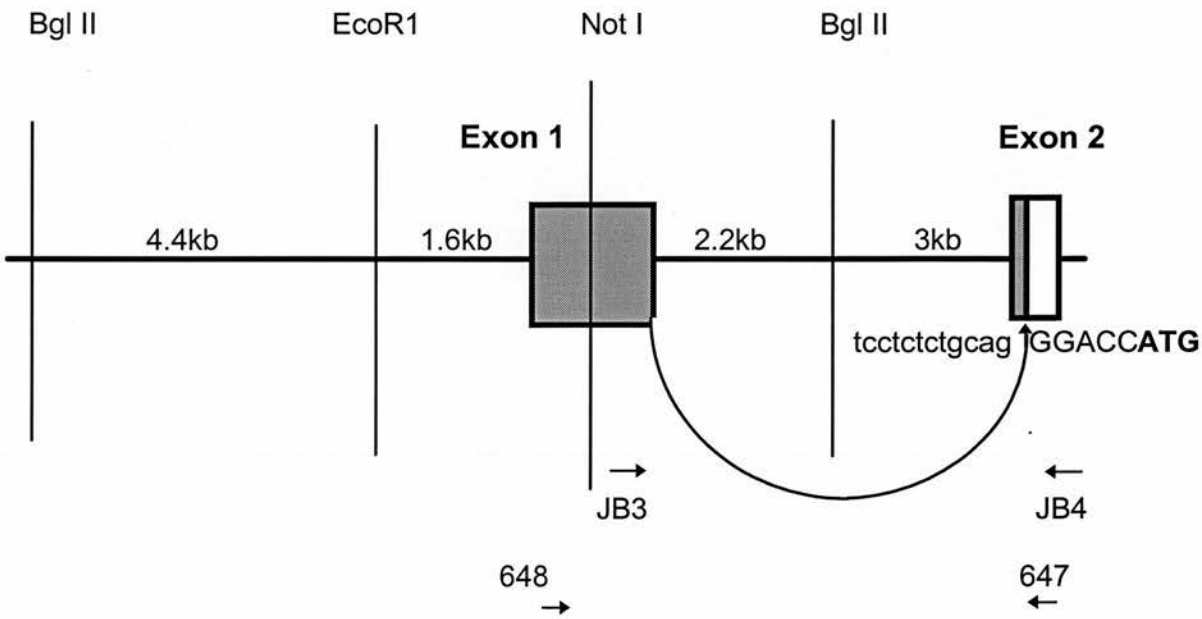


Figure 5.7 Partial restriction map of the 5' end of the ALK-1 gene. The position of exon 1 and 2 is given in relation to the Not1 site in exon 1, and the flanking Bgl II sites. There is also an EcoR1 site 1.6kb 5' of exon 1. The positions of primers JB3, JB4, 648 and 647 are marked. These primers were used as oligo-nucleotide probes for southern blotting and in PCR reactions to calculate the distances given.

The curved arrow marks the putative alternative splice between exon 1 and exon 2 as described in the text. The sequence at this splice junction is given with the start codon common to both variants marked in bold.

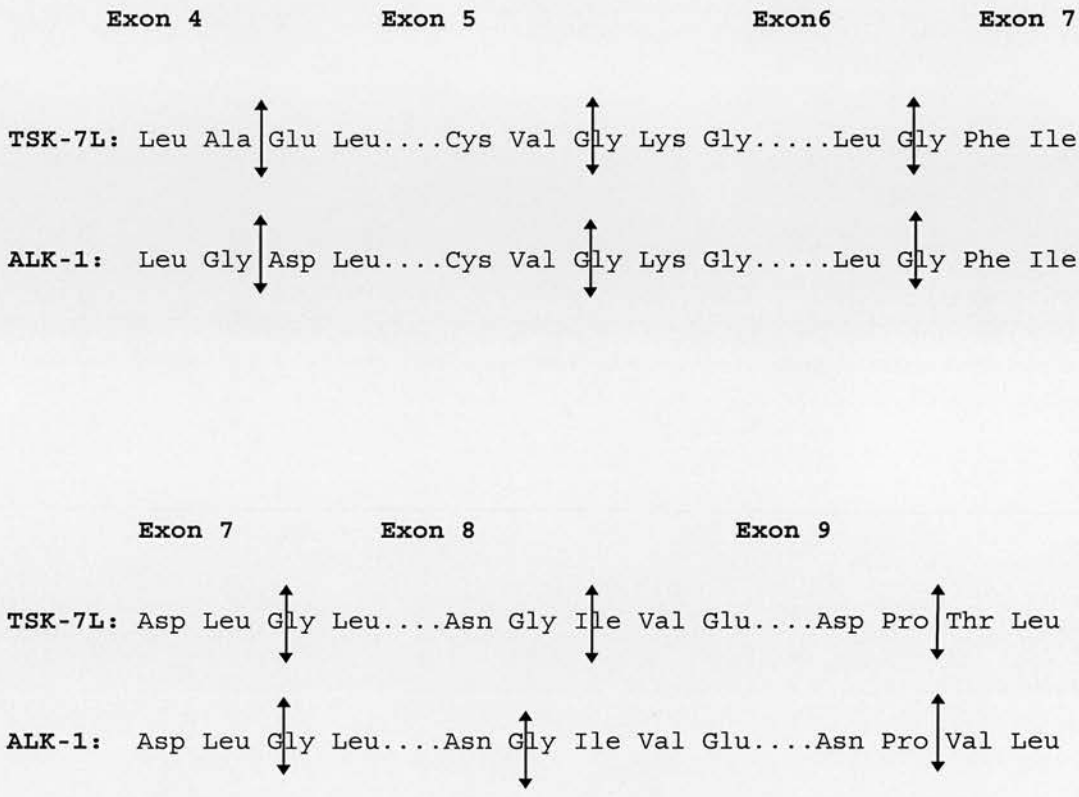


Figure 5.8. Comparison of the splice sites at the amino acid level between human ALK-1 and the mouse serine-threonine kinase type I receptor TSK7L. The comparison is confined to the kinase domain as there is little homology between the extracellular domains of the molecules.

Marks the position of each splice in the amino acid sequence.

As noted in the text, the splice sites between exons 5,6,7,8 and 9 all fall between 1st and second bases of a codon which encodes a neutral amino acid, most commonly glycine.

6. Mutation Analysis of the ALK-1 Gene.

Initial mutation analysis of the ALK-1 gene by rtPCR and sequencing from whole lymphocyte RNA was carried out by D. Johnson. Mutations in ALK-1 were identified in 3 families using this approach. These included a 3 base pair deletion identified in family 9, deleting a serine in kinase subdomain II downstream from the ATP-binding site. Mutations were also identified in families 33 and 40 (Johnson et.al. 1996) ascertained previously. Details of these mutations are included in **table 6.1** and **figure 6.1.**

A. Families Used for Mutation Analysis.

The ALK-1 gene was sequenced in 12 individuals. Six of these individuals came from HHT families that had evidence of linkage to 12q, either by a 2 point LOD score greater than 2 at $\theta = 0$ for a marker close to the ALK-1 region or formal exclusion of linkage to the endoglin region of chromosome 9q34. The remaining 6 individuals analysed were reported as having HHT, but came from families too small for linkage analysis, or for whom we only had a sample from a single affected individual.

Details of each linkage group are given below.

i. Definite 12 linked Families:

Family 2

A British family showing definite linkage to 12q34 as described in chapter 4. The family member chosen for sequencing had severe nosebleeds and gastro-intestinal haemorrhage as well as many classical telangiectases.

Family 3

A British family with HHT linked to 12q34 (Porteous et.al. 1994). An individual from this family had a recombination that was crucial in narrowing the candidate interval for HHT2.

Family 17

An American family from the Northern United States with 6 affected members, This family is described in (McAllister et.al. 1994b). Linkage to 12q was demonstrated by other workers.

ii. Families with Exclusion of Linkage to 9q34

Family 1

A large kindred previously published as showing definite exclusion of linkage to 9q34 on multipoint and 2 point linkage analysis (Porteous et.al. 1994). However as described previously, few markers in the ALK-1 region gave significantly raised LOD scores, but failed to exclude the region.

Family 5

This large family showed exclusion of 9q34 by multipoint analysis. However no markers from the ALK-1 region gave conclusive linkage or exclusion of linkage on two point analysis. As discussed in chapter 4, individuals in this family with nosebleeds and an affected first degree relative were erroneously diagnosed as affected.

Family 92

A large kindred from the Southern United States who have many members who might be mildly affected. This diagnostic confusion led to the use of only a small set of samples from a core kindred, allowing exclusion of 9q34 by other workers, but not providing sufficient data to confirm linkage to 12q34. The individual sequenced had frequent nosebleeds and unequivocal telangiectasis on the lips, ears and tongue.

iii. Individuals for Whom No Conclusive Linkage Data Was Available.

Individual A

An Individual from a small Dutch family. The individual was diagnosed as being affected by his own physician and has an affected sister.

Individual B

An individual assessed in person, with frequent nosebleeds and telangiectases. He thought that he might be a distant member of a large

Mormon family with HHT, who also come from the Salt-Lake City region. This family has been reported in an abstract as having a maximum 2 point LOD score of 2.4 for markers close to ALK-1, indicating possible linkage.

Individual C

A member of a very small British family who have no crossovers with markers from the ALK-1 region, but insufficient members are available to establish definite linkage.

Individual D

The only member of a large family from Texas who was assessed. They reported more than monthly nosebleeds with an affected parent. On examination by their own physician they had a few suspect red marks on the lips but no definite telangiectases. Two other family members had a history of pulmonary arteriovenous malformations.

Individual E

An American individual assessed by their own family physician with a history of severe nosebleeds and muco-cutaneous telangiectases. A member of a large family that has not been investigated in depth.

Individual F

A British individual, assessed in person with frequent nosebleeds and muco-cutaneous telangiectases. While the only member of this family from

whom we had a sample, there is an extensive history of other family members with nosebleeds, muco-cutaneous telangiectases and severe gastro-intestinal haemorrhage.

B. Identification of Nine Novel Mutations in the ALK-1 Gene.

For each sample, the entire coding region of the ALK-1 gene was amplified in 9 separate PCR reactions and sequenced using the methods described previously.

Nine mutations were identified in the ALK-1 gene. Of these, there were 2 amino acid changes in the extracellular domain, 3 created premature stop codons, one was a one base pair insertion and there were 3 amino acid changes in the kinase domain. The mutations found are summarised in **table 6.1,** and their position in the ALK-1 protein shown diagramatically in **figure 6.1.** For each mutation identified, a rapid assay was developed which enabled a population screen to be carried out on over 100 individuals. Details of these assays are also given in **table 6.1.**

Figure 6.2 shows an example of the family tree, sequence data generated and polymorphism assay for family 92.

Exon	Mutations Found (Position)	Effect on Protein	Assay for Mutation	No Of Normal Chromosomes Screened
III	G to T (150) G to A (200)	Tryptophan to Cysteine Arginine to Glutamine (in extracellular domain)	Removes ApaL1 site Sequencing	224 212
IV	G to A (423) G to T (475)	Intracellular Stop Intracellular Stop	Creates Afl II site Removes HinF I site	224 224
V	None			
VI	3 bp deletion (694)	Deletion of serine in kinase domain*		
VII	Ins T (865) C to A (925) G to T (998)	Frame Shift/stop Premature Stop Serine to Isoleucine (In kinase domain)	Heteroduplex Creates Stu I site Creates Alw I site	224 230 228
VIII	C to T (1120) T to G (1126) G to A (1232)	Arg to Trp Met to Arg * Arg to Glu* (In kinase domain)	Creates Dra III site	232
IX	C to A 1270	Proline to Threonine (In kinase domain)	Sequencing	232

Table 6.1
Table of the Mutations found with their position in the ALK-1 gene. As well as the 9 mutations described in the text, the 3 mutations described previously are included and marked *.

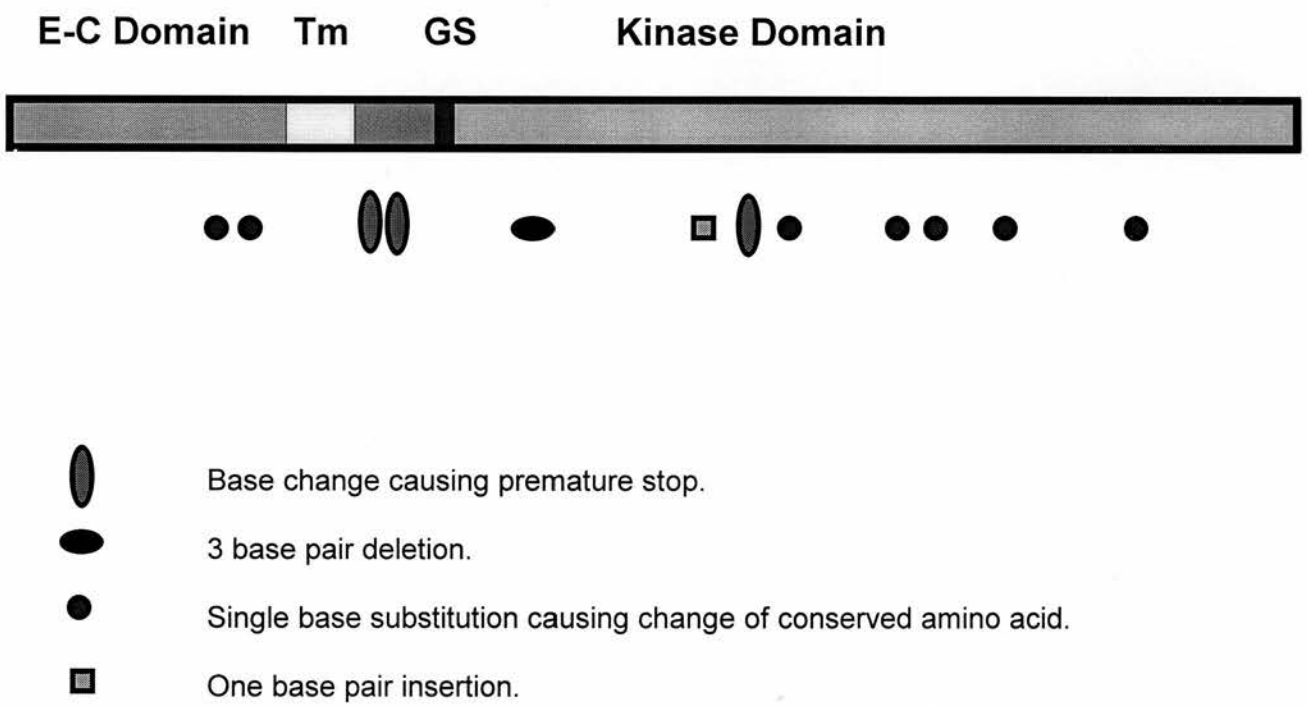


Figure 6.1 Diagram showing position and type of mutation in relation to domain of the ALK-1 molecule. **EC** Extracellular, **TM** Transmembrane, **GS** Glycine-serine domain.

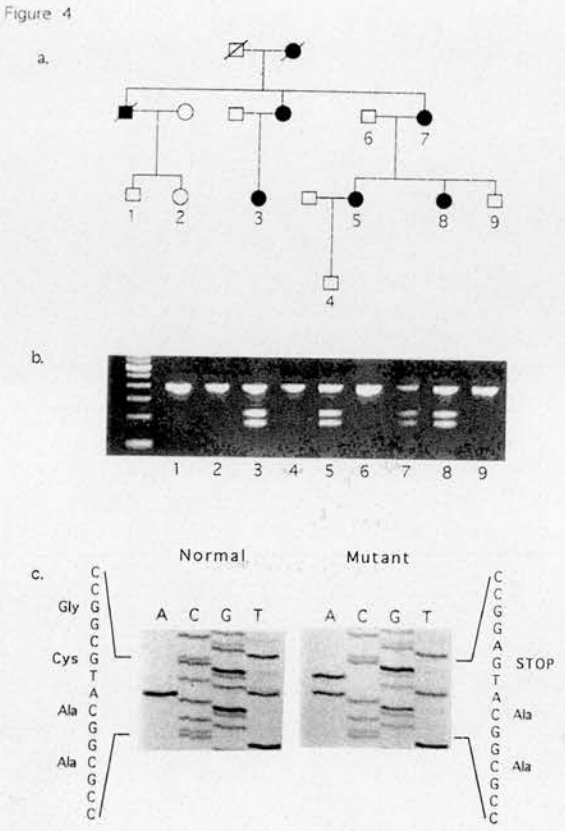


Figure 6.2. Example of data generated during mutation analysis.

a. Family tree of a section of family 92. (not ascertained in this study)

b. PCR assay for the mutation C925A identified in exon 7 in this study, which causes a premature stop. The sequence change creates a Stu1 site. All the affected individuals therefore show a wild-type PCR product of 350bp size as well as 2 digested bands where the mutated PCR product is cut by Stu1. The ladder in the left hand lane is Gibco-BRL 100bp ladder.

c. Sequence of exon 7 from subcloned PCR product showing the sequence of the normal and mutant alleles.

C. Confirmation of Significance of Mutations.

i. Polymorphism Screen.

For each of eight of the mutations described above, a panel of over 112 unrelated individuals was screened to ensure that the mutation found did not arise in the general population as a common polymorphism. The remaining mutation was not seen in 106 unnaffected individuals. The polymorphism panel comprised unaffected "married in" members of the HHT families. They were therefore unrelated, and, on the whole representative of the population from which the families came.

The chance of a single individual carrying an allele that occurs with an frequency (q) of 0.01, in a population in Hardy-Weinberg equilibrium is $2pq+q^2=0.02$ (assuming q^2 is negligible). The probability of the individual not having the polymorphism approximates to $1-2pq=0.98$. When screening 112 normal individuals from the same population the probability that all 112 will not carry the allele is $(0.98)^{112}=0.1$ or 10%. Screening 112 individuals would therefore give a 90% probability of detecting a population polymorphism with an allele frequency of 0.01 or greater.

ii. Significance of the Amino Acid Substitutions.

Where a mutation caused an amino acid change in ALK-1, a BLAST search was performed using the surrounding amino acid sequence to estimate the degree of conservation of the region that was changed. Six

published sequences were taken to represent the serine-threonine kinase receptor family from humans and other frequently studied species. Alignment was performed using the BLASTp program, and by eye. The results of this approach used to look at the 3 intracellular amino acid changes are given in **figure 6.3.** There was also conservation of the amino acid that was changed across many other gene family members.

A similar method used for the two extracellular mutations was less successful as the only proteins to show similarity in the extracellular domain are the rat and mouse homologues of ALK-1. However, in both these, the amino acid that was changed is conserved across species. This is shown in **figure 6.4**.

D. Efficiency of detection of Mutations

Table 6.2 shows the proportion of mutations detected in each of the three groups, described at the beginning of this chapter.

In all families with evidence of linkage to chromosome 12, or exclusion of linkage to 9q34, a mutation was found. Mutations were also found in 3 of the 6 individuals for whom no linkage data was available. Of the 3 individuals in whom no mutation was found, the individual D was incidentally re-assessed before results were released, by the physician who originally saw them. This individual was felt to be unnaffected. A

subsequent linkage study on this family also demonstrated linkage to 9q34 (D.Marchuk personal communication). Data from this individual was therefore not included in **table 6.2.**

Including the 3 mutations already identified, sequencing has therefore found 12 mutations in 14 affected individuals, representing every 12q34 linked family that has been identified , every family with definite exclusion of 9q34 and 3/5 individuals for whom no definite linkage data was available.

E. Identification of two polymorphisms within the ALK-1 Gene

Two polymorphisms were identified within the ALK-1 gene. The first, a single base change in exon 8 did not result in an amino acid change and was only seen once. The second was a common variant in the intron 3' of exon 3. The sequence is shown in **figure 6.5.** Of the 12 controls run, 6 were homozygous for the C, 4 were heterozygous and 2 were homozygous for the T. Of the two healthy volunteers who were homozygous for the T, both had had severe nosebleeds as children, requiring cauterisation. The significance of this in the normal population or as a modifying factor for symptoms of HHT remains uncertain.

Figure 6.3 Alignment either manual or by BLAST search of Amino Acid sequence surrounding each of the 3 amino-acid changes in the kinase domain of the ALK-1 gene. This shows conservation of the mutated amino acid across members of the same gene family in both humans and other frequently studied organisms, both vertebrate and invertebrate.

Where the one letter amino acid code is given there is exact alignment of the residue. "-" is used to denote a different amino acid in that position.

C1120T leading to Arginine to Tryptophan

Mutant Sequence	PRVGTK**W**YMAPE
Normal Sequence	PRVGTK**R**YMAPE
Human TGF-b type I receptor	-RVGTK**R**YMAPE
Human Activin type I receptor	-RVGTK**R**YMAPE
Mus Musculus TGF-b type I receptor	-RVGTK**R**YMAPE
Drosophila ATR-1 receptor	-RVGTK**R**YMAPE
Drosophila SAX	P-VGTK**R**YMAPE
C-Elegans C32D5.2 Receptor	--VGTK**R**YMAPE

G998T leading to Serine to Isoleucine

Mutant Sequence	IAHRDFK**I**RNVLVK
Normal Sequence	IAHRDFK**S**RNVLVK
Human TGF-b type I receptor	IAHRD-K**S**-N-LVK
Human Activin type I receptor	IAHRD-K**S**-N-LVK
Mus Musculus TGF-b type I receptor	IAHRD-K**S**-N-LVK
Drosophila ATR-1 receptor	---RD-K**S**-N-LVK
Drosophila SAX	-AHRD-K**S**-N-LV-
C-Elegans C32D5.2 Receptor	IAHRD-K**S**-N--VK

C1270A Leading to Proline to Threonine

Mutant Sequence
VEDYRP**T**FYDVVP
Normal Sequence
VEDYRP**P**FYDVVP

Human TGF-b type I receptor	-EDY--**P**-YD-VP
Human Activin type I receptor	-E-Y--**P**-YD-VP
Mus Musculus TGF-b type I receptor	-EDY--**P**-YD-VP
Drosophila ATR-1 receptor	---Y--**P**-YDVV-
Drosophila SAX	-E-Y--**P**FYDVVP
C-Elegans C32D5.2 Receptor	No Homology identified

Figure 6.4 Alignment by BLAST search of Amino Acid sequence surrounding each of the 2 extracellular amino-acid changes in the ALK-1 gene. The one letter amino acid code is given where there is exact alignment. "-" is used to denote a different amino-acid.

This figure shows conservation of the mutated amino acid in the known ALK-1 homologues in other species.

G150T leading to Tryptophan to Cysteine

Mutant Sequence	TCRGA**C**CTVVLV
Normal Sequence	TCRGA**W**CTVVLV
Rattus Norwegicus SKR3 Receptor	-C-G-**W**CTVVLV
Mus Musculus Activin IB receptor	-C-G-**W**CTVVLV

G200A Leading to Arginine to Glutamine

Mutant Sequence	HPQEH**Q**GCGNLH
Normal Sequence	HPQEH**R**GCGNLH
Rattus Norwegicus SKR3 Receptor	HPQ--**R**GCG-L-
Mus Musculus Activin IB receptor	HPQ--**R**GCG-L-

	Number Analysed	Number of Mutations Found
Families Linked to 12q	6	6
Linkage Excluded to 9q	3	3
No Linkage data Available	5	3
Total	**14**	**12**

Table 6.2 Showing the proportion of mutations found in the three groups, families known to link to chromosome 12q, families with exclusion of 9q34 (HHT1) on linkage, and individuals for whom no linkage data was available. All 12 known mutations are included in this table.

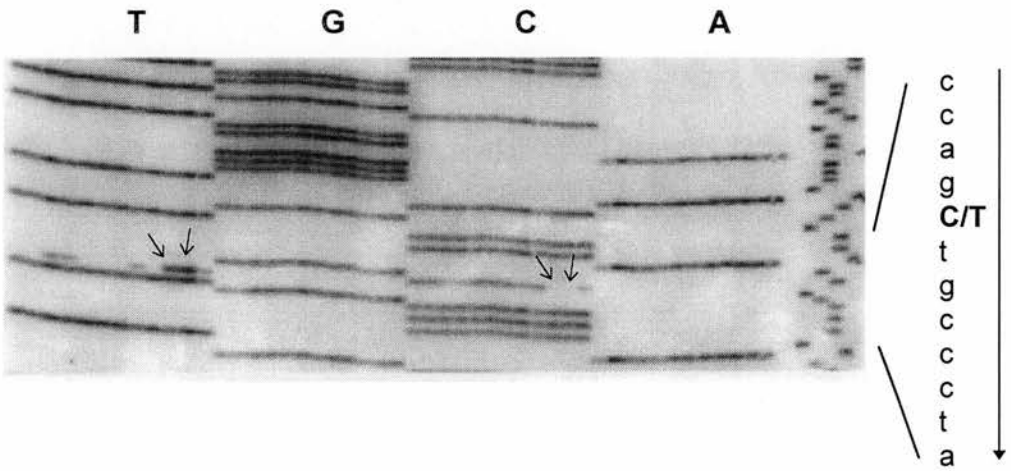


Figure 6.5. Sequencing reaction to show the polymorphism 3′ of exon 3. 12 individuals were sequenced. All 12 of each termination reaction were loaded adjacent to each other on the gel. The sequence is labelled in the orientation of the coding strand.

The arrows mark the sequencing lanes 10 and 11, from individuals who are homozygous for the **T** in the position shown, with no band seen in the "C" lane. Similarly, lanes 1,2,5,6,7 and 9 do not have a band in the "T" lane at this position, these individuals are homozygous for the "**C**" variant.

7. Identification of MAD Proteins Expressed in Endothelial Cells.

A. Design of Primers for degenerate PCR

Alignment of published amino acid sequences for different MAD proteins reveals domains of marked homology in the 5' and 3' ends of the protein with a variable region between. These domain have been named the MH1 and MH2 domains (Attisano personal communication). The sequence alignment used is given in **figure 7.1**. Primers were designed to these homologous domains to allow amplification of the variable central region. Although this would potentially necessitate the amplification of fragments up to 1.8kb in the largest identified MAD homologues, this also allows more accurate identification of the fragment as the 5' and 3' domains do not have sufficient diversity of sequence. **Table 7.1** shows the primer sequence chosen, corresponding amino acid sequence and level of degeneracy.

B. Amplification of a single fragment corresponding to DPC4.

An initial degenerate PCR reaction using cDNA from the ECRF24 cell line and conditions outlined in the materials and methods gave 2 bands around the 220 and 110 bp size. When cloned and sequenced, these were identified as RNA binding proteins L24 and L18. As they do not have exact

amino acid sequence homology to the primers used and are apparently not related to MAD proteins, this was taken as a non-significant result.
A subsequent PCR reaction using a more stringent "touchdown" program gave three fragment at low concentration. These fragments were of size 600bp, 800bp and 900bp, The bands were gel extracted and T/A cloned (see materials and methods). Only the 800bp fragment proved to be present at high enough concentration to allow cloning.

Sequencing of the several clones of this fragment from both ends identified the fragment as being derived from DPC4, a MAD homologue previously described as deleted in pancreatic carcinoma.

The PCR reaction was repeated using the EC1 and EC2 libraries (described previously) as template. PCR products corresponding to MADR1 and DPC4 were obtained.

5' Domain

MADR1.....110....EFP**F**GSKQKE**VC**I**NPYHY**K**RV**ES**P**VLPPVLVPRHSEYNP...

JV18-1.....150....EYA**F**NLKKDE**VC**V**NPYHY**Q**RV**ET**P**VLPPVLVPRHTEILT...

DPC-4.....118....QYA**F**DLKCDS**VC**.**NPYHY**E**RV**VS**P**GIDLSGLTLQSNAPS...

CE-Sma-2...108....RFCYESGQKDI**C**I**NPYHY**K**RV**HATGVLPPVLVPRYSEKP...

3' Domain

MADR1.....307....**rfclgllsnv**n**r**nsti**e**nt**rr**hig**k**gvhlyyvggevy...

JV18-1.....310....**rfclgllsnv**n**r**natv**e**mt**rr**hig**r**gvrlyyigevfa...

DPC-4.....361....**rfclgqlsnv**h**r**teai**e**ra**r**l**hig**k**gvql**eckgegdv...

CE-Sma-2.....259....kis**lg**lf**snv**n**r**nati**e**nt**rr**hig**n**gvkltyvrsngs...

Figure 7.1 Alignment of parts of the conserved 5' and 3' domains of MAD protein family members. Amino acid residues common to all 3 human molecules are given in bold. CE-Sma-2 is a MAD homologue from C.elegans. The sequences used to identify the best consensus sequence for degenerate PCR are underlined and are common to members of the gene family from widely divergent species.

5' Primer

Amino Acid Sequence	P	Y	H	Y	(K/E)	R	V
Nucleotide Sequence	CCU	UAU	CAU	UAU	AAA	AGA	GU
	C	C	C	C	G G	C C	
	A					A	
	G					G	

3' Primer

Amino Acid Sequence	H	I	G	K	G	V
Nucleotide sequence	CAU	AUU	GGU	AAA	GGU	GUU
	C	C	C	G	C	C
		A	A		A	A
			G		G	G

Table 7.1 Degenerate primer sequence derived from conserved domains of MAD proteins. The 5' primer has 1024 possible permutations, with a possible approximate TM in the range of 50 to 66. The 3' primer has 768 permutations with a TM range of 48 to 60.

8. Comparison of 9-Linked and 12-Linked Families

Table 8.1 shows a comparison of the incidence of PAVMs and gastro-intestinal haemorrhage between the chromosome 9 linked families 6 and 8 and the chromosome 12 linked families 5 and 9. Even within this small data set, the higher number of PAVMs in endoglin linked HHT compared to non chromosome 9 linked HHT is significant. There is no significant difference in incidence of gastro-intestinal haemorrhage.

This data set is expanded to include the incidence of PAVMs in all published families in **table 8.2.** From these tables it is evident that there is a much greater reported incidence of pulmonary arteriovenous malformations in families with linkage to chromosome 9, than those with linkage to chromosome 12.

	Affected Family members with PAVMs	Affected members without PAVMs	Total family members
Endoglin linked families (6 and 8)	4	10	14
Endoglin excluded families (5 and 9)	0	32	32
Total	4	42	46

P<0.01 (Fisher's exact test)

	Affected Family members with GI bleeding	Affected members without GI bleeding	Total family members
Endoglin linked families (6 and 8)	1	13	14
Endoglin excluded families (5 and 9)	6	26	32
Total	7	39	46

NS (Fisher's exact test)

Table 8.1 Comparison of incidence of PAVMs and Gastro-intestinal (GI) bleeding between families linked to chromosome 9 and endoglin (6 and 8) and families showing exclusion of this region on linkage analysis.(5 and 9). The incidence of PAVMs is higher in endoglin linked HHT, there is no significant difference in frequency of GI bleeding between the two groups.

	Affected Family members with PAVMs	Affected members without PAVMs	Total family members
Endoglin linked families [a]	37	91	128
Endoglin excluded families [b]	1	87	88
Total	38	178	216

$p<0.001$ by χ^2 analysis

a: Data from families 01,02,32,56 (McAllister et.al., families F and A (Shovlin et.al.) as well as families 6 and 8.

b: Data from families 1,2,3,4 (Porteous et.al.), 17 and 33 (McAllister et.al.) and family T (Shovlin et.al.), as well as families 5 and 8.

Table 8.2. Comparison of PAVM frequency between endoglin linked and endoglin excluded families, using all available data from published as well as newly ascertained families. The very low frequency of PAVMs in the non endoglin linked HHT families is highly significant.

Chapter 9. Discussion

A.Clinical Features of Hereditary Haemorrhagic Telangiectasia.

i. Clinical Differences between HHT1 and HHT2

Data from patients were, on the whole, obtained during short visits to large family groups. Data collection from the American families was further complicated by lack of permission to contact the patient's own medical practitioner, preventing detailed examination of case notes for confirmation of all symptoms reported. In this situation it was not possible to request any detailed screening tests as a part of the research project.

All conclusions are, therefore, based on reported symptoms and simple screening tests available during the family visits.

Patients with mutations in either Endoglin or ALK-1 were prone to severe epistaxis and transfusion dependant gastro-intestinal haemorrhage. The incidence of gastro-intestinal haemorrhage was the same in Endoglin and ALK-1 patients.

However, in families with mutations in the endoglin gene, or those with established linkage to chromosome 9, the incidence of pulmonary arteriovenous malformations (PAVMs) was statistically much higher than in those with ALK-1 mutations.

The higher incidence of PAVMs in those with endoglin mutations remains to be explained. To date none of the British or American families with known ALK-1 mutations, either contacted in this study, or the earlier one by Porteous et.al. (1992) have evidence of a pulmonary arteriovenous malformation. There have been preliminary reports of two patients with PAVMs in a large Mormon kindred with a known ALK-1 mutation. (Macdonald personal communication).

There are clinical implications for patient management arising from this finding. People with endoglin mutations are at a definite high risk of having assymptomatic PAVMs and developing the associated complications. They represent a particularly high risk subset of HHT patients and should be carefully screened with an established effective technique for detecting PAVMs. Patients with a family history of PAVMs are more likely to have endoglin mutations and should also be considered to be at high risk. Current evidence would suggest that chest x-ray and arterial blood gas measurement or pulmonary arteriovenous shunt measurement are the best modalities for clinical screening.

Clinical screening for patients with ALK-1 mutations is a more complex issue. Given the very small number of patients with known ALK-1 mutations and PAVMs, none of whom have been published, patients with ALK-1 mutations can be reassured that their risk of having a PAVM is low. Currently it is still sensible to advocate simple screening techniques such

as chest x-ray and pulse oximetry for patients with ALK-1 mutations, until this situation is clarified.

There have been several reports of CAVMs clustering within families and occurring in association with PAVMs. Family 6 described in this text has two living individuals with PAVMs, one of whom had a symptomatic CAVM which required neuro-surgery. In the same family, an individual who died shortly before they were assessed had both PAVMs and a CAVM.

As yet, the case for screening for CAVMs in HHT is unproven, as it remains uncertain whether pre-symptomatic screening or treatment of CAVMs is beneficial.

iii Potential Diagnostic Tests.

While many people with HHT have florid symptoms and present no diagnostic difficulty, particularly at a young age, patients can show few symptoms.

Mutation analysis from genomic DNA is now possible both for Endoglin (McAllister et.al. 1995, Marchuk 1997) and ALK-1 (data presented above).

An early diagnostic test will allow pre-symptomatic testing of individuals. In those who are unaffected, it will allow them to be reassured that they are not at risk of developing symptoms, and they will not need screening for

PAVMs and CAVMs. In the future, with increased understanding of the disease, it may be possible to initiate treatment before symptoms develop, or at least to avoid factors which may be shown to cause deterioration of skin, gut and nose lesions. Presence of an Endoglin mutation in an asymptomatic patient would justify screening for PAVMs.

As discussed above, being able to differentiate between HHT1 and HHT2 is of some benefit in sparing patients screening investigations that are both expensive and uncomfortable. Further studies are necessary to establish whether this can usefully be applied to a clinical setting. In the future, assays of endoglin protein expression by monocytes may be useful in replacing DNA analysis.

B How Mutation in the ALK-1 Gene Leads to Disease.

i.The effect of mutation on the ALK-1 gene expression and function

All the mutations reported so far in the ALK-1 gene leave the transmembrane domain of the protein intact,10 mutations disrupt the kinase domain, either by frame shift, premature stop or altering a highly conserved amino-acid. Two of the mutations would be expected to alter important amino acids in the extracellular domain, but do not cause a frame shift.

It is possible that all the mutations lead to a null allele for ALK-1 or a null allele equivalent, a version of the protein which has no biological action. The initial studies of ALK-1 using rtPCR from whole blood RNA showed the mutant allele in some reactions but not others. However as expression of ALK-1 in blood is probably at a very low level, rtPCR is not a reliable indicator of ALK-1 expression in endothelial cells.

Alternatively, a defective ALK-1 protein product that is necessary for the disease pathogenesis might be produced. There are, therefore, three possible hypotheses for the way mutation in ALK-1 affects the endothelial cell. These will be dealt with in turn.

a.Haploinsufficiency for ALK-1

The absence of one functioning ALK-1 allele in endothelial cells may be sufficient to cause the disease. Cells with a lower level of ALK-1 expression than normal might be more susceptible to an insult which irreversibly triggers the formation of a telangiectasis. Such an insult could be an external stimulus with subsequent failure of a normal repair mechanism due to insufficient ALK-1 signalling, or the level of ALK-1 expression might drop below a particular threshold, causing susceptibility to hydrostatic pressure in the vessel. This mechanism is consistent with the location of telangiectases which arise more or less at random, but in very specific locations. The mouth, ears, nose and fingers are particularly exposed to heat, cold and trauma, whereas the trunk is less vulnerable.

b. A Second Genetic Hit at the ALK-1 Locus

A second somatic mutation leads to loss of the wild type ALK-1 allele in an endothelial cell with subsequent formation of a telangiectasis. The genetic insult may arise during foetal life, but the lesion will remain dormant until a further external stimulus causes it to enlarge. However the microscopic studies of the skin lesions, showing dilatation of the terminal venule as the first visible lesion, rather than a proliferative event makes this hypothesis seem less likely.

However, in adult polycystic kidney disease type 1 (APKD-1), the epithelial cells lining cysts have been shown to have loss of the functioning copy of the APKD-1 gene. (Brassier and Henske 1997, Quian et.al. 1996). The cysts are not usually present in the kidney at birth, but appear in varying numbers during childhood and adult life. Their enlargement as cysts is not a neoplastic process. Similarly, in HHT, loss of a functioning copy of ALK-1 in endothelium during embryological development may predispose a region to the formation of vascular malformations, without predisposing to cell proliferation.

c. A Dominant Negative Effect on Receptor Function.

Expression of the mutant ALK-1 protein might interfere either with the activity of the normal ALK-1 protein, or with another signalling molecule, causing a greater than 50% reduction in effective ALK-1 signalling within the cell. For this to be the case, the component of the signalling complex

has to be present in limiting concentrations. This could be either ALK-1 itself, the type II receptor, ligand or the downstream signalling molecule. In favour of this hypothesis, several workers have shown that the overexpression of a non functioning type I receptor can act in a dominant negative fashion (Feng et.al. 1995, Yamomoto et.al. 1996). However, these experiments were performed in cultured cell systems and are unlikely to accurately demonstrate the true effect in vivo. One study has shown a low TGF-β receptor type II level in endothelial cells dividing in 3 dimensional culture, which has decreased from when the cells are grown as monolayers, suggesting that TGF-β receptor type II concentration may be a control mechanism and may indeed be at a limiting concentration (Sankaar et.al. 1996). This model, however fails to explain how discreet lesions occur at random locations.

The mechanism by which endoglin mutation causes HHT also remains unclear, haploinsufficiency, two hit and dominant negative hypotheses have all been suggested, although recent discovery of several true null mutations has made the dominant negative hypothesis seem unlikely for this gene. (Shovlin et.al. 1997, Marchuk et.al. personal communication).

C. The Interaction Between Endoglin and ALK-1

I. Interaction at the cellular level.

The discovery that mutations in both endoglin and ALK-1 could cause a very similar phenotype suggested an interaction between these two molecules in the endothelial cell. There are three possible ways that this interaction could occur, either individually or in combination. These are shown diagramatically in **figure 9.1**.

a.Endoglin interacts directly with ALK-1 in a signalling complex.

The presence of endoglin on the endothelial cell surface facilitates signalling by ALK-1(**Figure 9.1a)**. This need not be an exclusive interaction, and endoglin may be involved with the TGF-β receptor complex as well. However there is, as yet, no direct evidence of a direct interaction between ALK-1 and Endoglin, but a protein molecule of the size of ALK-1 has been shown to co-precipitate when endoglin is precipitated with a monoclonal antibody.

B. Endoglin changes local concentration of TGF-β without involvement in signalling.

Endoglin acts as a TGF-β binding protein in its own right without apparent signalling activity. It has been shown to be present in serum and extracellular matrix, where it has presumably been released from the endothelial cell

surface. In a similar situation, betaglycan has been shown to sequester TGF-β2 in the extracellular matrix, possibly acting as a slow release mechanism.

Lower endoglin concentrations in extracellular matrix might be responsible for higher peak activity of TGF-β signalling through the ALK-1 receptor, followed by rapid fall-off in signal. **(Figure 9.1b)**

c.Endoglin acts through a separate signalling complex to ALK-1.

Endoglin interacts with a different signalling complex to ALK-1, but that the pathways of these two signalling complexes interact within the cell **(Figure 9.1c)**.

ii. Why Endoglin mutations cause a higher incidence of PAVMs

The reason why endoglin mutation predisposes individuals to developing PAVMs remains unclear. Recent evidence also suggests that patients with endoglin mutations have earlier onset of symptoms (Porteous et.al. 1997). This implies one of three possibilities.

a. Endoglin has an effect on other pathways in addition to ALK-1 signalling.

The expression of endoglin in cell types other than endothelial cells, and the interaction between endoglin and the TGF-β receptor type I suggest that this is the case. It may or may not be significant in the pathogenesis of the disease.

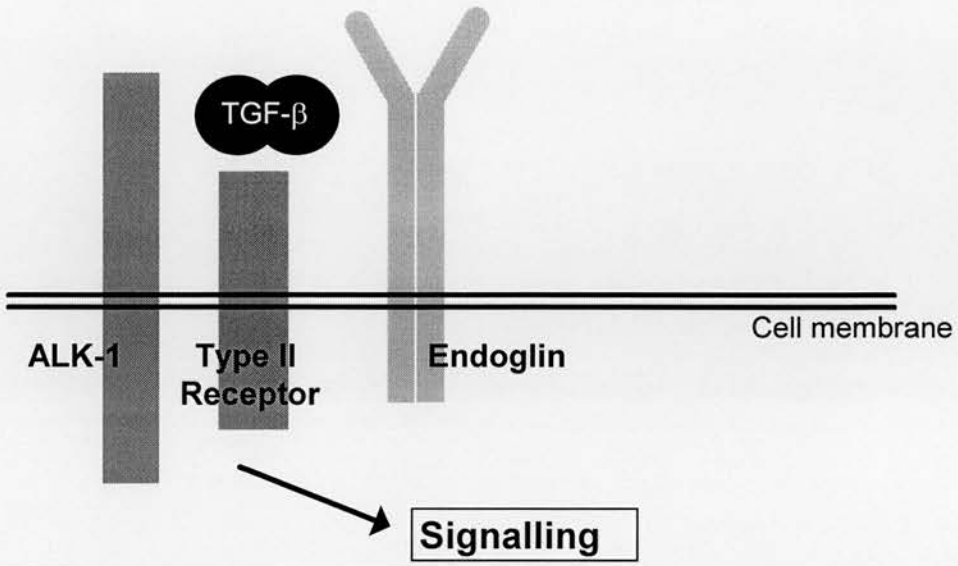


Figure 9.1a. The first hypothesis that endoglin interracts directly with the ALK-1 signalling complex.

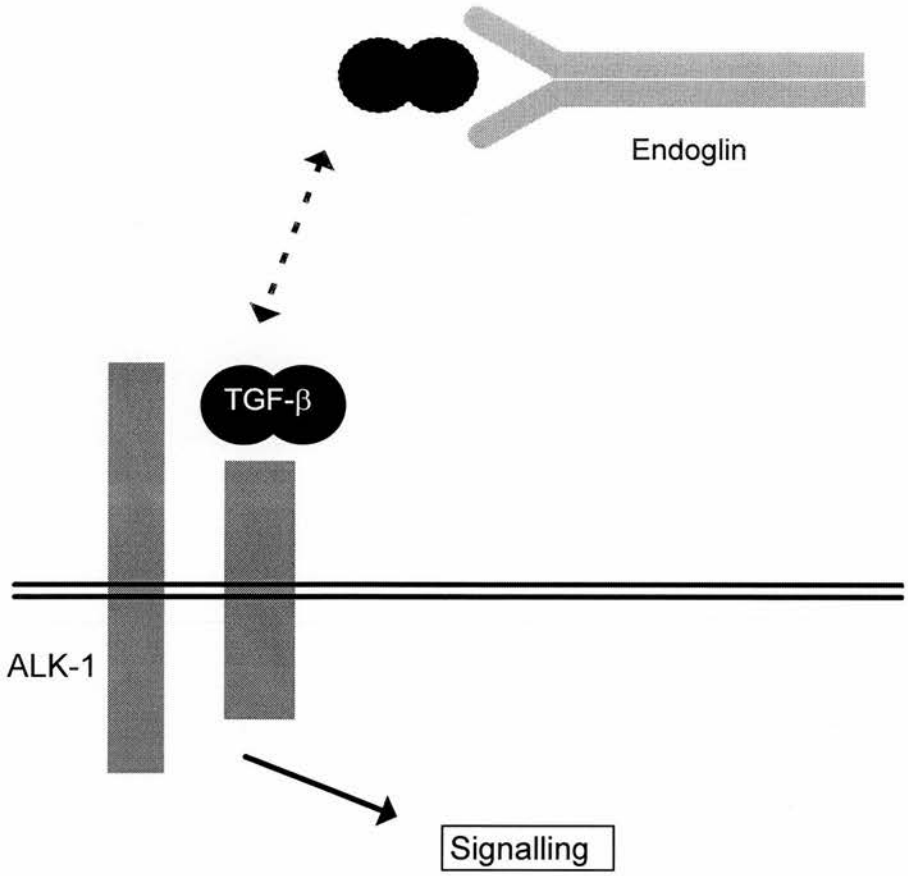


Figure 9.1b. Endoglin, possibly located in the extracellular matrix, acts as a reservoir for TGF-β, releasing it to allow signalling through the ALK-1 signalling complex. Alternatively, the presence of endoglin on the cell surface might increase the concentration of TGF-β close to the cell surface without interacting directly with the receptor complex.

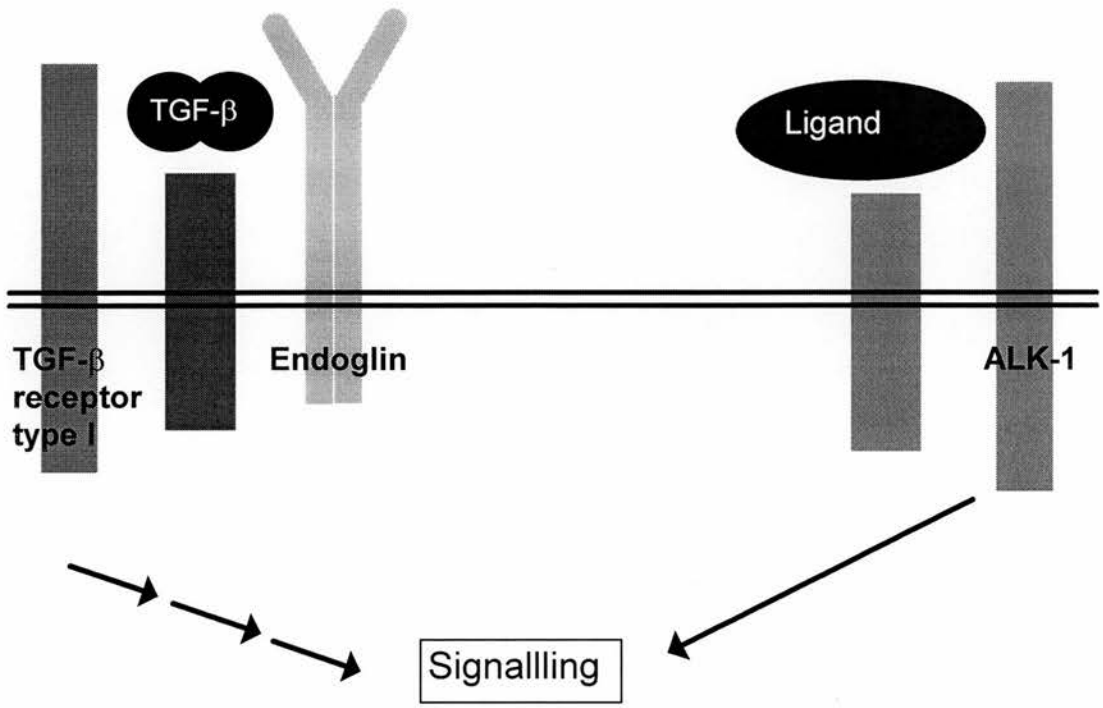


Figure 9.1c. Endoglin functions in conjunction with the TGF-β receptor complex. ALK-1 signals independantly (with either the same or a different ligand and type II receptor), the signalling pathways interact within the cell.

b. Inability of the endothelial cell to compensate for endoglin deficiency.

The endothelial cell may be less able to compensate for endoglin deficiency than ALK-1 deficiency. Endoglin may have a unique role in the endothelial cell, whereas ALK-1 is known to have several homologues expressed in the same cell type. These include the activin and TGF-β receptor type I. One of these may be partially active in the same signalling pathways as ALK-1. Loss of ALK-1 signalling may therefore be partially compensated by these receptors. This conservation may be a important in early lung development and similarly may make the enlargement of pre-formed lung lesions to clinically detectable size less likely in the case of ALK-1 deficiency.

c. Endoglin is important earlier in fetal development.

Endoglin might have an important role in early development of fetal vasculature that is not shared by ALK-1, particularly in the lung. The possibility that PAVMs are abnormalities of fetal development is not a new idea (Wirth et.al. 1996), enlargement of lesions to detectable size could occur with age and hormonal change.

iii A new model for ALK-1 and Endoglin interaction in HHT causation.

Given the small amount of knowledge, any model for the interaction between endoglin and ALK-1 is purely speculative. Such a model has to be designed to take into account the known facts of endoglin interaction with the TGF-β receptor complex (Lastres et.al. 1996) with apparent

modification of TGF-β signalling in the monocyte, as well as the degree of shared phenotype between endoglin and ALK-1 mutation. The existence of other receptors, signalling pathways and downstream interacting proteins in the cell cannot be ignored.

At the cell surface, it could be hypothesised that endoglin enhances ALK-1 signalling, but inhibits aspects of TGF-β signalling through the TGF-β receptor type I. Increased expression of endoglin in an endothelial cell would decrease TGF-β receptor type I signalling, but increase ALK-1 signalling.

If the ligand for ALK-1 is TGF-β, this would lead to a balance between TGF-β and ALK-1 signalling in the cell, which would in part be controlled by level of endoglin expression **(Figure 9.2).** Loss of ALK-1 signalling would therefore reduce ALK-1 signalling alone without direct effect on the TGF-β receptor. Reduction in endoglin would not only cause reduced ALK-1 signalling, but also increased TGF-β signalling. There would also be an effect of endoglin mutation in situations where ALK-1 was not expressed, but TGF-β receptor I signalling was important.

Further experimentation, determining the true ligand for ALK-1 and effect of ALK-1 signalling on endothelial cell growth in conjunction with endoglin are necessary to give evidence for or against this model.

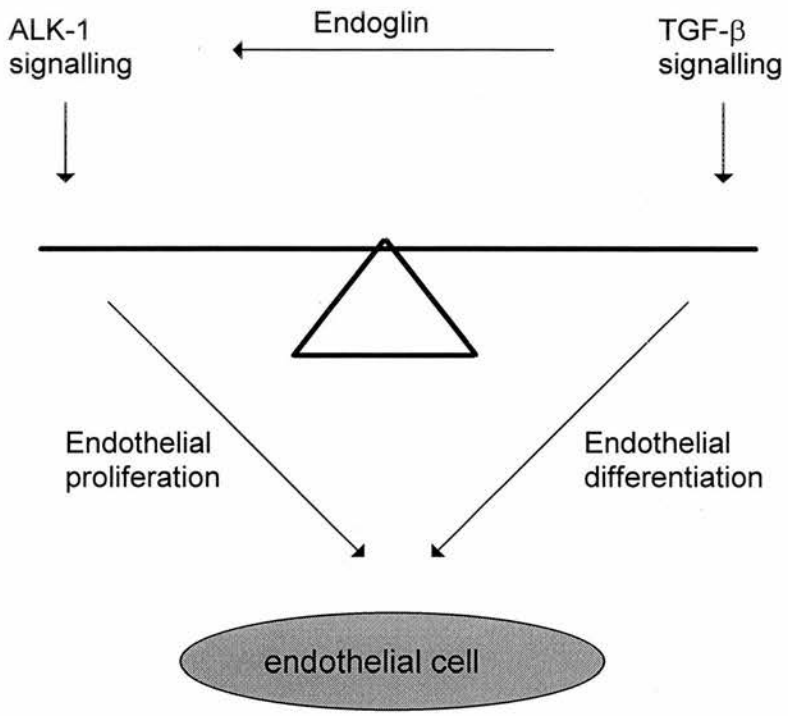


Figure 9.2. Model for the role of ALK-1 and endoglin. The expression of endoglin increases signalling through the ALK-1 signalling complex, and reduces signalling through the TGF-β complex. This would alter the balance between endothelial cell differentiation and endothelial cell proliferation.

D. The Genomic Structure of ALK-1

i.Potential start sites and promoters of ALK-1 transcription

The two published sequences for the full length ALK-1 cDNA are identical in the coding region, but have different alternative 5′ untranslated regions. The sequences diverge 7 base pairs upstream from the putative start codon.

When a comparison is made between the genomic structure and the published sequences, the sequence published by Attisano (Attisano et.al. 1993) begins in exon 2, reading through the transcription start site which is also in exon 2. The sequence published by ten Djike et.al. starts in exon 1 and splices to a consensus splice junction 7bp upstream from the start codon. Both published sequences were determined by 5′ race, and it is possible that the sequence published by Attisano is the result of genomic contamination in one of the RACE reactions.

Alternatively, there may be two alternative transcription start sites for ALK-1, implying the existence of two different promoter regions. This might be a control mechanism, with the two different promoters responsive to different factors, either in the same type of cell, or in different tissues.

A large (approx 4.4kb) transcript as well as a 2.2 kb transcript is identified on northern blotting by using the extracellular domain of ALK-1 as probe.

This could be an alternative 5′ untranslated region or an alternative 3′ untranslated region, which may have a regulatory mechanism. The size differences between the published ALK-1 cDNA sequences are not sufficient to explain the different transcript sizes, although the published sequences may not be full length. Alternatively the large transcript may represent an alternative coding region or an undiscovered homologue of ALK-1.

ii Splice positions within the ALK-1 gene.

None of the ALK-1 gene exons are exceptionally large or widely spaced. The coding region of the gene spans less than 20kb of genomic DNA with exon 1 lying 5.5kb upstream. A comparison of the splice junctions between intracellular domains of ALK-1 and the mouse putative TGF-β receptor type I TSK-7L. showed homology within the kinase domain, but marked differences in the extracellular domain. The position of introns in both ALK-1 and TSK-7L is almost identical between exons 4 and 10. All four splice sites between exons 5,6,7,8 and 9 fall between 1st and second bases of a codon encoding a neutral amino acid. This is the case in both human ALK-1 and mouse TSK-7L. This may be a result of the common ancestry of the intracellular domains, but the degree of conservation suggests a possible function. The positioning of splice sites would allow alternatively spliced products of ALK-1 to be formed which would remain in the same frame, but would be expected to have variant kinase activity. However, at present, there is no current evidence to suggest that alternate splicing is

physiologically important. Alternative splicing of a different nature has been shown to occur with other members of the serine-threonine kinase receptor family in both humans and Drosophila. (Attisano et.al. 1992)

The conservation of position of introns within the kinase domain of two related type I serine threonine kinase receptors from different species, indicates that this is likely to be conserved in many other related molecules, facilitating the determination of genomic structure of other type I serine-threonine kinase receptors.

iii Non-Consensus splice sequence 5′ of exon 6

The splice sequence 5′ to exon 6 varies marginally from the expected consensus. Rather than the expected donor splice consensus CRG/gt$^{a}_{g}$ag the sequence at this site is TAG/gcaag. While this is unusual, a computerised survey of such non-consensus splice donors (Senapathy et.al. 1990) has shown the TAG/gcaag to be the most common variant.

The function of this variant splice sequence is unknown, it may be of no functional significance. The recently described variant consensus AT-AC has been shown to require its own family of proteins for splicing (Tarn and Steitz 1996), adding a specific splicing control mechanism. However, there is no reason to suspect that this is the case for the G/gcaag splice variant. If the splice variant at this site has a function, it may affect the proportion of correctly spliced ALK-1 message containing exon 6 that is formed.

E. Degenerate rtPCR revealing the presence of DPC4 in endothelial cells.

The only members of the MAD protein family identified by degenerate rtPCR from endothelial cell RNA were DPC4 and MADR1.

As the DPC4 was successfully amplified in a single round of PCR, using primers to amplify from the 5′ conserved domain to the 3′conserved domain, it seems likely that DPC4 is expressed at significant levels in endothelial cells. The absence of other MAD protein homologues in the reaction is most likely to be due to one of several factors which enhanced DPC4 amplification preferentially. The reaction conditions used may have favoured DPC4 amplification, or the DPC4 may have been in higher concentration and therefore preferentially amplified. Alternatively the other homologues may have had a variation in sequence at the primer sites, not taken into account during primer design.

It is not known whether the ALK-1 signalling pathway itself involves DPC4. DPC4 (OMIM *600993) was first identified as a tumour suppresser gene deleted in pancreatic carcinoma. It seems unlikely that DPC4 gene mutations would cause HHT, as a phenotype involving other tissues, including susceptibility to various forms of malignancy would be anticipated.

Chapter 10. Conclusions.

Clinical Features of HHT

At the clinical level, this study has shown that despite diagnostic difficulties, due to the frequency of nosebleeds in the general population, clinical diagnostic criteria can be drawn up which if fulfilled enable a positive diagnosis to be made. There will still remain a number of patients, usually at a younger age for whom diagnosis is difficult on clinical grounds.

It has also been conclusively demonstrated that patients with endoglin mutations have a much higher rate of pulmonary arteriovenous malformations and therefore of associated complications. In this study, no patients came to light who had ALK-1 mutations and a pulmonary arteriovenous malformation. It is unknown whether PAVMs will remain exclusively associated with endoglin mutations. There is a subjective impression that cerebral arteriovenous malformations are more commonly associated with endoglin mutation, but this remains to be confirmed. It has not, so far, been possible to draw conclusions about other features of the disease.

The presence of an endoglin mutation in a family, or a family history of PAVM is therefore an indication for early screening of all family members for asymptomatic PAVMs. Further work is required to say whether patients with ALK-1 mutations are genuinely "safe" from the point of view

of PAVMs and CAVMs, or should be screened, but in a less aggressive manner.

Molecular genetics of HHT.

The detection of a large number of mutations in ALK-1 in patients with HHT confirms that this is the HHT2 locus. The additional fact that all the large families we have studied either show linkage to chromosome 9q34, chromosome 12q13, or have mutations in ALK-1 (on 12q13) suggests that involvement of any third locus for HHT is likely to be rare, possibly with a distinct phenotype. The Italian family reported with a high incidence of liver involvement may be an example of such a family.

The method for genomic sequencing of ALK-1 proved very successful (detecting mutations in all of the chromosome 12q34 linked families studied) and is applicable in a clinical setting. The low incidence of PAVMs in this group makes discovery of a mutation reassuring for the individual affected, although ideally, a good laboratory test for endoglin mutations would also be clinically useful.

The nature of mutations found in combination with *in vitro* data on other serine-threonine kinase receptors is compatible with the hypothesis that expression of a mutant ALK-1 protein is necessary for disease to occur. However there is currently no direct evidence to favour this dominant

negative hypothesis over the hypothesis that haploinsufficiency for ALK-1 is sufficient to cause disease. The implications for novel therapeutic approaches to HHT include pharmacological means of raising ALK-1 expression or signalling in endothelial cells.

ALK-1 receptor function.

The function of ALK-1 remains unknown, although its identification as the HHT2 gene suggests that it has an important role in vascular endothelium. This could be either in controlling the response of differentiated vessels to injury, in maintaining the structure of differentiated small vessels, or in controlling angiogenesis and blood vessel formation.

The fact that mutations in both ALK-1 and Endoglin cause a similar disease suggests that they both have a role in the same endothelial cell pathway. The nature of this interaction, either directly at the cell surface, or intracellularly remains to be determined.

The true ligand and intracellular signalling pathway for ALK-1 remain unknown. The demonstration of DPC4 expression in the endothelial cell is the first report of this, but mainly serves to underline that these molecules are involved in signalling in many cell types. It is not known from this study whether DPC4 is involved in ALK-1 signalling.

Future Research.

With the availability of a gene-based test for ALK-1 mutations, the potential clinical impact of this needs to be assessed in a formal trial. For endoglin mutations, molecular testing requires further refinement, and in the long term, endoglin expression levels by monocyte derived cell lines may prove a quicker and more effective test.

The ability to differentiate at the molecular level between HHT1 and HHT2 will enable further study to identify the clinical differences between the two. This will allow future targeting of screening on those at risk of specific complications. Identification of the genetic factors that act as modifier loci for HHT, for example, pre-disposing to development of PAVMs may prove possible in the future. As well as providing further insight into the genetic basis of HHT, this may also allow identification of individuals at particular risk.

Study towards further understanding of how ALK-1 and Endoglin interact and signal will be important, both for understanding signalling in the endothelial cell, as well as identifying future means of treatment for HHT. Further investigation of the promoter region may identify factors which enable up-regulation of ALK-1 expression as a possible means of treatment for HHT. Identification of promoter elements which control endothelial specific gene expression may prove important in future gene-

therapy approaches aiming to introduce cell-specifically expressed genes into cells.

The expression pattern of ALK-1 remains to be fully determined. As well as the tissue distribution, it is unknown whether ALK-1 is expressed at different stages of the cell cycle, and what effect ALK-1 signalling has on the cell.

The construction of knockout mice for endoglin and ALK-1 is the next step likely to give significant insight into how these genes function and cause disease. They may also prove to be valuable models for novel therapeutic approaches.

Understanding the control of blood vessel growth and differentiation may prove important in the treatment of disorders other than HHT. Many common disorders involve abnormal blood vessel development or function. Diabetic retinopathy and retinopathy of prematurity involve the development of fragile blood vessels at the back of the eye. All forms of tumour require the recruitment of blood vessels in order to grow. Control of blood vessel growth is, therefore, one potential therapeutic target in the treatment of many common forms of disease.

References

AAssar, O.S., Friedman, C.M. and White, R.I., Jr. (1991) The natural history of epistaxis in hereditary hemorrhagic telangiectasia. Laryngoscope 101, 977-980.

Ainsworth, P.J., Surh, L.C. and Coulter-Mackie, M.B. (1991) Diagnostic single strand conformational polymorphism, (SSCP): a simplified non-radioisotopic method as applied to a Tay-Sachs B1 variant. Nucleic Acids Research 19, 405-406.

Attisano, L., Wrana, J.L., Cheifetz, S. and Massague, J. (1992) Novel activin receptors: distinct genes and alternative mRNA splicing generate a repertoire of serine/threonine kinase receptors. Cell 68, 97-108.

Attisano, L., Carcamo, J., Ventura, F., Weis, F.M., Massague, J. and Wrana, J.L. (1993) Identification of human activin and TGF beta type I receptors that form heteromeric kinase complexes with type II receptors. Cell 75, 671-680.

Attisano, L., Wrana, J.L., Lopez-Casillas, F. and Massague, J. (1994) TGF-beta receptors and actions. [Review] [105 refs]. Biochimica. et. Biophysica. Acta 1222, 71-80.

Attisano, L., Wrana, J.L., Montalvo, E. and Massague, J. (1996) Activation of signalling by the activin receptor complex. Molecular. &. Cellular. Biology. 16, 1066-1073.

Attisano, L. and Wrana, J.L. (1996) Signal transduction by members of the transforming growth factor-beta superfamily. [Review] [127 refs]. Cytokine. and. Growth Factor. Reviews. 7, 327-339.

Baarends, W.M., van Helmond, M.J., Post, M., van der Schoot, P.J., Hoogerbrugge, J.W., de Winter, J.P., Uilenbroek, J.T., Karels, B., Wilming, L.G. and Meijers, J.H. (1994) A novel member of the transmembrane serine/threonine kinase receptor family is specifically expressed in the gonads and in mesenchymal cells adjacent to the mullerian duct. Development 120, 189-197.

Bauer, T., Britton, P., Lomas, D., Wight, D.G., Friend, P.J. and Alexander, G.J. (1995) Liver transplantation for hepatic arteriovenous malformation in hereditary haemorrhagic telangiectasia. Journal. of. Hepatology 22, 586-590.

Bean, W.B. (1971) Retinal involvement in hereditary hemorrhagic telangiectasia. Archives. of. Ophthalmology. 86, 726

Berg, J.N., Guttmacher, A.E., Marchuk, D.A. and Porteous, M.E. (1996) Clinical heterogeneity in hereditary haemorrhagic telangiectasia: are pulmonary

arteriovenous malformations more common in families linked to endoglin?. Journal. of. Medical. Genetics 33, 256-257.

Bernard, G., Mion, F., Henry, L., Plauchu, H. and Paliard, P. (1993) Hepatic involvement in hereditary hemorrhagic telangiectasia: clinical, radiological, and hemodynamic studies of 11 cases. Gastroenterology 105, 482-487.

Bideau, A., Plauchu, H., Jacquard, A., Robert, J.M. and Desjardins, B. (1980) [Genetic aspects of Rendu-Osler disease in Haut-Jura: convergence of methodological approaches of historic demography and medical genetics]. [French]. Journal. de. Genetique. Humaine. 28, 127-147.

Bideau, A., Brunet, G., Heyer, E., Plauchu, H. and Robert, J.M. (1992) An abnormal concentration of cases of Rendu-Osler disease in the Valserine valley of the French Jura: a genealogical and demographic study. Annals. of. Human. Biology. 19, 233-247.

Bird, R M, Hammarsten, J F, Marshall, R A, and Robinson, R R. (1957) A family re-union: a study of hereditary haemorrhagic telangiectasia. New England Journal of Medicine 257, 105-109.

Bonyadi, M., Rusholme, S.A., Cousins, F.M., Su, H.C., Biron, C.A. and Farrall, M. (1997) Mapping of a major genetic modifier of embryonic lethality in TGF beta 1 knockout mice. Nature Genetics 15, 207-211.

Bown, S.G., Swain, C.P., Storey, D.W., Collins, C., Matthewson, K., Salmon, P.R. and Clark, C.G. (1985) Endoscopic laser treatment of vascular anomalies of the upper gastrointestinal tract. Gut 26, 1338-1348.

Brant, A.M., Schachat, A.P. and White, R.I. (1989) Ocular manifestations in hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber disease). [Review] [32 refs]. American. Journal. of. Ophthalmology. 107, 642-646.

Brasier, J.L. and Henske, E.P. (1997) Loss of the polycystic kidney disease (PKD1) region of chromosome 16p13 in renal cyst cells supports a loss-of-function model for cyst pathogenesis. Journal of Clinical Investigation 99, 194-199.

Braverman, I.M., Keh, A. and Jacobson, B.S. (1990) Ultrastructure and three-dimensional organization of the telangiectases of hereditary hemorrhagic telangiectasia. Journal. of. Investigative. Dermatology. 95, 422-427.

Buscarini, E., Buscarini, L., Danesino, C., Piantanida, M., Civardi, G., Quaretti, P., Rossi, S., Di Stasi, M. and Silva, M. (1997) Hepatic vascular malformations in hereditary hemorrhagic telangiectasia: Doppler sonographic screening in a large family. Journal. of. Hepatology 26, 111-118.

Carcamo, J., Weis, F.M., Ventura, F., Wieser, R., Wrana, J.L., Attisano, L. and Massague, J. (1994) Type I receptors specify growth-inhibitory and transcriptional responses to transforming growth factor beta and activin. Molecular. &. Cellular. Biology. 14, 3810-3821.

Carcamo, J., Zentella, A. and Massague, J. (1995) Disruption of transforming growth factor beta signaling by a mutation that prevents transphosphorylation within the receptor complex. Molecular. &. Cellular. Biology. 15, 1573-1581.

Cheifetz, S., Weatherbee, J.A., Tsang, M.L., Anderson J.K., Mole J.E., Lucas R and Massague,J. (1987) The transforming growth factor-beta system, a complex pattern of cross-reactive ligands and receptors. Cell 48, 409-15.

Cheifetz, S., Hernandez, H., Laiho, M., ten Dijke, P., Iwata, K.K. and Massague, J. (1990) Distinct transforming growth factor-beta (TGF-beta) receptor subsets as determinants of cellular responsiveness to three TGF-beta isoforms. Journal. of. Biological. Chemistry. 265, 20533-20538.

Cheifetz, S., Bellon, T., Cales, C., Vera, S., Bernabeu, C., Massague, J. and Letarte, M. (1992) Endoglin is a component of the transforming growth factor-beta receptor system in human endothelial cells. Journal. of. Biological. Chemistry. 267, 19027-19030.

Chen, Y., Lebrun, J.J. and Vale, W. (1996) Regulation of transforming growth factor beta- and activin-induced transcription by mammalian Mad proteins. Proceedings of the National Academy of Sciences of the United States of America 93, 12992-12997.

Cotton, R.G., Rodrigues, N.R. and Campbell, R.D. (1988) Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutations. Proceedings of the National Academy of Sciences of the United States of America 85, 4397-4401.

Davis, D.G. and Smith, J.L. (1971) Retinal involvement in hereditary hemorrhagic telangiectasia. Archives. of. Ophthalmology. 85, 618-21 passim.

Dickson, M.C., Martin, J.S., Cousins, F.M., Kulkarni, A.B., Karlsson, S. and AkhurstRJ. (1995) Defective haematopoiesis and vasculogenesis in transforming growth factor-beta 1 knock out mice. Development 121, 1845-1854.

Dutton, J.A., Jackson, J.E., Hughes, J.M., Whyte, M.K., Peters, A.M., Ussov, W. and Allison, D.J. (1995) Pulmonary arteriovenous malformations: results of treatment with coil embolization in 53 patients. AJR. American. Journal. of. Roentgenology. 165, 1119-1125.

Ebner, R., Chen, R.H., Lawler, S., Zioncheck, T. and Derynck, R. (1993) Determination of type I receptor specificity by the type II receptors for TGF-beta or activin. Science 262, 900-902.

Feng, X.H., Filvaroff, E.H. and Derynck, R. (1995) Transforming growth factor-beta (TGF-beta)-induced down-regulation of cyclin A expression requires a functional TGF-beta receptor complex. Characterization of chimeric and truncated type I and type II receptors. Journal of Biological Chemistry 270, 24237-24245.

Ference, B.A., Shannon, T.M., White, R.I., Jr., Zawin, M. and Burdge, C.M. (1994) Life-threatening pulmonary hemorrhage with pulmonary arteriovenous malformations and hereditary hemorrhagic telangiectasia. Chest 106, 1387-1390.

Fernandez-Ruiz, E., St-Jacques, S., Bellon, T., Letarte, M. and Bernabeu, C. (1993) Assignment of the human endoglin gene (END) to 9q34-->qter. Cytogenetics. &. Cell Genetics 64, 204-207.

Folkman, J. and D'Amore, P.A. (1996) Blood vessel formation: what is its molecular basis? [comment]. [Review] [20 refs]. Cell 87, 1153-1155.

Forker, E L, and Bean, W B. (1963) Retinal arteriovenous aneurysm in hereditary haemorrhagic telangiectasia. Arch. Intern. Med. 3, 778-783.

Franzen, P., ten Dijke, P., Ichijo, H., Yamashita, H., Schulz, P., Heldin, C.H. and Miyazono, K. (1993) Cloning of a TGF beta type I receptor that forms a heteromeric complex with the TGF beta type II receptor. Cell 75, 681-692.

Franzen, P., Heldin, C.H. and Miyazono, K. (1995) The GS domain of the transforming growth factor-beta type I receptor is important in signal transduction. Biochemical. &. Biophysical. Research. Communications. 207, 682-689.

Ginsburg D., Handin R., Bonthron D., Donlon T., Bruns G., Latt S., and Orkin S. (1985) Human von Willebrand factor (vWF) : isolation of complementary DNA (cDNA) clones and chromosomal localisation. Science 228, 1401-6

Gougos, A., St Jacques, S., Greaves, A., O'Connell, P.J., d'Apice, A.J., Buhring, H.J., Bernabeu, C., van Mourik, J.A. and Letarte, M. (1992) Identification of distinct epitopes of endoglin, an RGD-containing glycoprotein of endothelial cells, leukemic cells, and syncytiotrophoblasts. International. Immunology 4, 83-92.

Gougos, A. and Letarte, M. (1990) Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. Journal. of. Biological. Chemistry. 265, 8361-8364.

Graff, J.M., Bansal, A. and Melton, D.A. (1996) Xenopus Mad proteins transduce distinct subsets of signals for the TGF beta superfamily. Cell 85, 479-487.

Guttmacher, A.E., McKinnon, W.C. and Upton, M.D. (1994) Hereditary hemorrhagic telangiectasia: a disorder in search of the genetics community [letter]. American. Journal. of. Medical. Genetics 52, 252-253.

Guttmacher, A.E., Marchuk, D.A. and White, R.I., Jr. (1995) Hereditary hemorrhagic telangiectasia [see comments]. [Review] [53 refs]. New England. Journal. of. Medicine 333, 918-924.

Gyapay, G., Morissette, J., Vignal, A., Dib, C., Fizames, C., Millasseau, P., Bernardi, G., Lathrop, M. and Weissenbach, J. (1994) The 1993-94 Genethon human genetic linkage map [see comments]. Nature Genetics 7, 246-339.

Haitjema, T., Disch, F., Overtoom, T.T., Westermann, C.J. and Lammers, J.W. (1995) Screening family members of patients with hereditary hemorrhagic telangiectasia. American. Journal. of. Medicine 99, 519-524.

Haitjema, T., Balder, W., Disch, F.J. and Westermann, C.J. (1996a) Epistaxis in hereditary haemorrhagic telangiectasia. Rhinology. 34, 176-178.

Haitjema, T., ten Berg, J.M., Overtoom, T.T., Ernst, J.M. and Westermann, C.J. (1996b) Unusual complications after embolization of a pulmonary arteriovenous malformation. Chest 109, 1401-1404.

Haitjema, T.J., Overtoom, T.T., Westermann, C.J. and Lammers, J.W. (1995) Embolisation of pulmonary arteriovenous malformations: results and follow up in 32 patients [see comments]. Thorax 50, 719-723.

Hanes, F M (1909) Multiple hereditary telangiectasis causing haemorrhage (hereditary haemorrhagic telangiectasia). Bull. Johns Hopkins Hospital 20, 63-73.

Harrison, D.F. (1970) Hereditary haemorrhagic telangiectasia and oral contraceptives. Lancet 1, 721

Harrison, D.F. (1982) Use of estrogen in treatment of familial hemorrhagic telangiectasia. Laryngoscope 92, 314-320.

Hashimoto, K. and Pritzker, M.S. (1972) Hereditary hemorrhagic telangiectasia. An electron microscopic study. Oral Surgery,. Oral Medicine,. Oral Pathology. 34, 751-768.

Heutink, P., Haitjema, T., Breedveld, G.J., Janssen, B., Sandkuijl, L.A., Bontekoe, C.J., Westerman, C.J. and Oostra, B.A. (1994) Linkage of hereditary haemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. Journal. of. Medical. Genetics 31, 933-936.

Hodgson, C H, Burchell, H B, Good, C A, and Clagett, O T. (1959) hereditary haemorrhagic telangiectasia and pulmonary arteriovenous fistula. New England Journal of Medicine 261, 625-635.

Hoodless, P.A., Haerry, T., Abdollah, S., Stapleton, M., O'Connor, M.B., Attisano, L. and Wrana, J.L. (1996) MADR1, a MAD-related protein that functions in BMP2 signaling pathways. Cell 85, 489-500.

Horn, G., Rabb, M.F. and Lewicky, A.O. (1981) Retinal telangiectasis of the macula: a review and differential diagnosis. [Review] [19 refs]. International. Ophthalmology. Clinics. 21, 139-155.

Hughes, J.M. (1994) Intrapulmonary shunts: coils to transplantation. [Review] [39 refs]. Journal. of. the. Royal. College. of. Physicians. of. London. 28, 247-253.

Jackson, J.E., Whyte, M.K., Allison, D.J. and Hughes, J.M. (1990) Coil embolization of pulmonary arteriovenous malformations. Cor. et. Vasa. 32, 191-196.

Jahnke, V. (1970) Ultrastructure of hereditary telangiectasia. Archives. of. Otolaryngology. 91, 262-265.

Johnson, D.W., Berg, J.N., Gallione, C.J., McAllister, K.A., Warner, J.P., Helmbold, E.A., Markel, D.S., Jackson, C.E., Porteous, M.E. and Marchuk, D.A. (1995) A second locus for hereditary hemorrhagic telangiectasia maps to chromosome 12. Genome Research. 5, 21-28.

Johnson, D.W., Berg, J.N., Baldwin, M.A., Gallione, C.J., Marondel, I., Yoon, S.J., Stenzel, T.T., Speer, M., Pericak-Vance, M.A., Diamond, A., Guttmacher, A.E., Jackson, C.E., Attisano, L., Kucherlapati, R., Porteous, M.E. and Marchuk, D.A. (1996) Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. Nature Genetics 13, 189-195.

Kaartinen, V., Voncken, J.W., Shuler, C., Warburton, D., Bu, D. and Heisterkamp, N. (1995) Abnormal lung development and cleft palate in mice lacking TGF-beta 3 indicates defects of epithelial-mesenchymal interaction. Nature Genetics 11, 415-421.

Kadoya, C., Momota, Y., Ikegami, Y., Urasaki, E., Wada, S. and Yokota, A. (1994) Central nervous system arteriovenous malformations with hereditary hemorrhagic telangiectasia: report of a family with three cases. Surgical. Neurology 42, 234-239.

Kawabata, M., Imamura, T., Miyazono, K., Engel, M.E. and Moses, H.L. (1995) Interaction of the transforming growth factor-beta type I receptor with farnesyl-protein transferase-alpha. Journal of Biological Chemistry 270, 29628-29631.

Keen, J., Lester, D., Inglehearn, C., Curtis, A. and Bhattacharya, S. (1991) Rapid detection of single base mismatches as heteroduplexes on Hydrolink gels. Trends in Genetics 7, 5

Kikuchi, K., Kowada, M. and Sasajima, H. (1994) Vascular malformations of the brain in hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber disease). Surgical Neurology 41, 374-380.

King, C.R., Lovrien, E.W. and Reiss, J. (1977) Central nervous system arteriovenous malformations in multiple generations of a family with hereditary hemorrhagic telangiectasia. Clinical Genetics 12, 372-381.

Kjeldsen, A.D., Vase, P. and Oxhoj, H. (1996) Hereditary hemorrhagic telangiectasia [letter; comment]. New England. Journal. of. Medicine 334, 331-332.

Krohel, G.B., Duke, M.A. and Mehu, M. (1987) Bloody tears associated with familial telangiectasis. Case report. Archives. of. Ophthalmology. 105, 1489-1490.

Lastres, P., Bellon, T., Cabanas, C., Sanchez-Madrid, F., Acevedo, A., Gougos, A., Letarte, M. and Bernabeu, C. (1992) Regulated expression on human macrophages of endoglin, an Arg-Gly-Asp-containing surface antigen. European. Journal. of. Immunology 22, 393-397.

Lastres, P., Martin-Perez, J., Langa, C. and Bernabeu, C. (1994) Phosphorylation of the human-transforming-growth-factor-beta-binding protein endoglin. Biochemical. Journal. 301, 765-768.

Lastres, P., Letamendia, A., Zhang, H., Rius, C., Almendro, N., Raab, U., Lopez, L.A., Langa, C., Fabra, A., Letarte, M. and Bernabeu, C. (1996) Endoglin modulates cellular responses to TGF-beta 1. Journal. of. Cell Biology. 133, 1109-1121.

Lesser, B.A., Wendt, D., Miks, V.M. and Norum, R.A. (1986) Identical twins with hereditary hemorrhagic telangiectasia concordant for cerebrovascular arteriovenous malformations. American. Journal. of. Medicine 81, 931-934.

Letsou, A., Arora, K., Wrana, J.L., Simin, K., Twombly, V., Jamal, J., Hoffmann, F.M., Gelbart, W.M., Massague, J. and et al. (1995) Drosophila Dpp signaling is mediated by the punt gene product: a dual ligand-binding type II receptor of the TGF beta receptor family. Cell 80, 899-908.

Liu, F., Hata, A., Baker, J.C., Doody, J., Carcamo, J., Harland, R.M. and Massague, J. (1996) A human Mad protein acting as a BMP-regulated transcriptional activator [see comments]. Nature 381, 620-623.

Lopez-Casillas, F., Cheifetz, S., Doody, J., Andres, J.L., Lane, W.S. and Massague, J. (1991) Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF-beta receptor system. Cell 67, 785-795.

Lopez-Casillas, F., Wrana, J.L. and Massague, J. (1993) Betaglycan presents ligand to the TGF beta signaling receptor. Cell 73, 1435-1444.

Lopez-Casillas, F., Payne, H.M., Andres, J.L. and Massague, J. (1994) Betaglycan can act as a dual modulator of TGF-beta access to signaling receptors: mapping of ligand binding and GAG attachment sites. Journal. of. Cell Biology. 124, 557-568.

Madri, J.A., Pratt, B.M. and Tucker, A.M. (1988) Phenotypic modulation of endothelial cells by transforming growth factor-beta depends upon the composition and organization of the extracellular matrix. Journal of Cell Biology 106, 1375-1384.

Madri, J.A., Bell, L. and Merwin, J.R. (1992) Modulation of vascular cell behavior by transforming growth factors beta. [Review] [37 refs]. Molecular Reproduction & Development 32, 121-126.

Mandriota, S.J., Menoud, P.A. and Pepper, M.S. (1996) Transforming growth factor beta 1 down-regulates vascular endothelial growth factor receptor 2/flk-1 expression in vascular endothelial cells. Journal of Biological Chemistry 271, 11500-11505.

Martin, J.S., Dickson, M.C., Cousins, F.M., Kulkarni, A.B., Karlsson, S. and Akhurst RJ. (1995) Analysis of homozygous TGF beta 1 null mouse embryos demonstrates defects in yolk sac vasculogenesis and hematopoiesis. [Review] [31 refs]. Annals of the New York Academy of Sciences 752, 300-308.

Martini, G.A. (1978) The liver in hereditary haemorrhagic teleangiectasia: an inborn error of vascular structure with multiple manifestations: a reappraisal. Gut 19, 531-537.

Massague, J. (1996) TGFbeta signaling: receptors, transducers, and Mad proteins. [Review] [20 refs]. Cell 85, 947-950.

Massague, J and Weiss-Garcia, F Serine/Threonine Kinase Receptors: Mediators of Transforming Growth Factor Beta Family Signals. [Review] Cancer Surveys 27, (*Cell Signalling)* 41-64.

McAllister, K.A., Grogg, K.M., Johnson, D.W., Gallione, C.J., Baldwin, M.A., Jackson, C.E., Helmbold, E.A., Markel, D.S., McKinnon, W.C. and Murrell, J. (1994a) Endoglin, a TGF-beta binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. Nature Genetics 8, 345-351.

McAllister, K.A., Lennon, F., Bowles-Biesecker, B., McKinnon, W.C., Helmbold, E.A., Markel, D.S., Jackson, C.E., Guttmacher, A.E., Pericak-Vance, M.A. and Marchuk, D.A. (1994b) Genetic heterogeneity in hereditary haemorrhagic telangiectasia: possible correlation with clinical phenotype. Journal. of. Medical. Genetics 31, 927-932.

McAllister, K.A., Baldwin, M.A., Thukkani, A.K., Gallione, C.J., Berg, J.N., Porteous, M.E., Guttmacher, A.E. and Marchuk, D.A. (1995) Six novel mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type 1 suggest a dominant-negative effect of receptor function. Human. Molecular. Genetics 4, 1983-1985.

McDonald, M.T., Papenberg, K.A., Ghosh, S., Glatfelter, A.A., Biesecker, B.B., Helmbold, E.A., Markel, D.S., Zolotor, A., McKinnon, W.C. and Vanderstoep, J.L. (1994) A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. Nature Genetics 6, 197-204.

Menefee, M.G., Flessa, H.C., Glueck, H.I. and Hogg, S.P. (1975) Hereditary hemorrhagic telangiectasia (Osler-Weber-Rendu disease). An electron microscopic study of the vascular lesions before and after therapy with hormones. Archives. of. Otolaryngology. 101, 246-251.

Merwin, J.R., Anderson, J.M., Kocher, O., Van Itallie, C.M. and Madri, J.A. (1990) Transforming growth factor beta 1 modulates extracellular matrix organization and cell-cell junctional complex formation during in vitro angiogenesis. Journal of Cellular Physiology 142, 117-128.

Merwin, J.R., Newman, W., Beall, L.D., Tucker, A. and Madri, J. (1991) Vascular cells respond differentially to transforming growth factors beta 1 and beta 2 in vitro. American Journal of Pathology 138, 37-51.

Mustoe, T.A., Pierce, G.F., Thomason, A., Gramates, P., Sporn, M.B. and Deuel, T.F. (1987) Accelerated healing of incisional wounds in rats induced by transforming growth factor-beta. Science 237, 1333-1336.

Naveau, S, Aubert A, Poynard T and Chaput J C. (1990) Long-term results of treatment of vascular malformations of the gastrointestinal tract by neodymium YAG laser photocoagulation. Digestive Diseases & Sciences.35,821-6

Niehrs, C. (1996) Growth factors. Mad connection to the nucleus [news; comment]. Nature 381, 561-562.

O'Connell, P.J., McKenzie, A., Fisicaro, N., Rockman, S.P., Pearse, M.J. and d'Apice, A.J. (1992) Endoglin: a 180-kD endothelial cell and macrophage restricted differentiation molecule. Clinical. &. Experimental. Immunology 90, 154-159.

Osler (1901) On a family form of recurring epistaxis associated with multiple telangiectases of the skin and mucous membranes. Bull. Johns Hopkins Hosp. 12, 333-337

Ott, J. (1991) Analysis of human genetic linkage, 2nd ed. Baltimore, Johns Hopkins University press.

Piantanida, M., Buscarini, E., Dellavecchia, C., Minelli, A., Rossi, A., Buscarini, L. and Danesino, C. (1996) Hereditary haemorrhagic telangiectasia with extensive liver involvement is not caused by either HHT1 or HHT2. Journal. of. Medical. Genetics 33, 441-443.

Plauchu, H., de Chadarevian, J.P., Bideau, A. and Robert, J.M. (1989) Age-related clinical profile of hereditary hemorrhagic telangiectasia in an epidemiologically recruited population. American Journal of Medical Genetics 32, 291-297.

Porteous, M.E., Burn, J. and Proctor, S.J. (1992) Hereditary haemorrhagic telangiectasia: a clinical analysis. Journal. of. Medical. Genetics 29, 527-530.

Porteous, M.E., Curtis, A., Williams, O., Marchuk, D., Bhattacharya, S.S. and Burn, J. (1994) Genetic heterogeneity in hereditary haemorrhagic telangiectasia. Journal. of. Medical. Genetics 31, 925-926.

Proetzel, G., Pawlowski, S.A., Wiles, M.V., Yin, M., Boivin, G.P., Howles, P.N., Ferguson, M.W. and Doetschman, T. (1995) Transforming growth factor-beta 3 is required for secondary palate fusion. Nature Genetics 11, 409-414.

Qian, F., Watnick, T.J., Onuchic, L.F. and Germino, G.G. (1996) The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type I. Cell 87, 979-987.

Quitt M, Froom P, Veisler A, Falber V, Sova J and Aghai E. (1990) The effect of desmopressin on massive gastrointestinal bleeding in hereditary haemorrhagic telangiectasia unresponsive to treatment with cryoprecipitate. Arch.Intern.Medicine 150,1744-1746.

Rendu (1896) Epistxis repetees chez un sujet porteurs de petits angiomes cutanes et muqeux. Bull.Mem.Soc.Med.Hop. 13, 731

Riggins, G.J., Thiagalingam, S., Rozenblum, E., Weinstein, C.L., Kern, S.E., Willson, J.K., Markowitz, S.D., Kinzler, K.W. and Vogelstein, B. (1996) Mad-related genes in the human. Nature Genetics 13, 347-349.

Roberts, A.B., Sporn, M.B., Assoian, R.K., Smith, J.M., Roche, N.S., Wakefield, L.M., Liotta, L.A., Falanga, V., Kehrl, J.H. and et al. (1986) Transforming growth factor type beta: rapid induction of fibrosis and angiogenesis in vivo and stimulation of collagen formation in vitro. Proceedings of the National Academy of Sciences of the United States of America 83, 4167-4171.

Sambrook, J. Fritsch, E F, and Maniatis T (1989) Molecular Cloning, a laboratory manual. Cold Spring Harbour Laboratory Press.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. Proceedings of the National Academy of Sciences of the United States of America 74, 5463-5467.

Sankar, S., Mahooti-Brooks, N., Bensen, L., McCarthy, T.L., Centrella, M. and Madri, J.A. (1996) Modulation of transforming growth factor beta receptor levels on microvascular endothelial cells during in vitro angiogenesis. Journal. of. Clinical. Investigation. 97, 1436-1446.

Schneider, M.D. and Brand, T. (1995) Molecular analysis of TGF beta signal transduction. Dominant-inhibitory mutations of the type II and type I TGF beta receptor. [Review] [41 refs]. Annals. of. the. New York. Academy. of. Sciences. 752, 309-316.

Schwartz, D.C. and Cantor, C.R. (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. Cell 37, 67-75.

Senapathy, P, Shapiro, M B, and Harris N L. (1990) Splice junctions, branch point sites, and exons: sequence statistics, identification, and application to genome project. Methods Enzymol 183, 252-278.

Shovlin, C.L., Hughes, J.M., Tuddenham, E.G., Temperley, I., Perembelon, Y.F., Scott, J., Seidman, C.E. and Seidman, J.G. (1994) A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. Nature Genetics 6, 205-209.

Shovlin, C.L., Winstock, A.R., Peters, A.M., Jackson, J.E. and Hughes, J.M. (1995) Medical complications of pregnancy in hereditary haemorrhagic telangiectasia. QJM. 88, 879-887.

Shovlin, C.L., Hughes, J.M., Scott, J., Seidman, C.E. and Seidman, J.G. (1997) Characterization of endoglin and identification of novel mutations in hereditary hemorrhagic telangiectasia. American Journal of Human Genetics 61, 68-79.

Southern, E.M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology 98, 503-517.

St-Jacques, S., Cymerman, U., Pece, N. and Letarte, M. (1994a) Molecular characterization and in situ localization of murine endoglin reveal that it is a transforming growth factor-beta binding protein of endothelial and stromal cells. Endocrinology 134, 2645-2657.

St-Jacques, S., Forte, M., Lye, S.J. and Letarte, M. (1994b) Localization of endoglin, a transforming growth factor-beta binding protein, and of CD44 and integrins in placenta during the first trimester of pregnancy. Biology. of. Reproduction. 51, 405-413.

Steele, J.G., Nath, P.U., Burn, J. and Porteous, M.E. (1993) An association between migrainous aura and hereditary haemorrhagic telangiectasia. Headache. 33, 145-148.

Sutton, H G. (1864) Epistaxis as an indication of impaired nutrition and of degeneration of the vascular system. Medical Mirror 1, 769.

Swinburne, A.J., Fedullo, A.J., Gangemi, R. and Mijangos, J.A. (1986) Hereditary telangiectasia and multiple pulmonary arteriovenous fistulas. Clinical deterioration during pregnancy. Chest 89, 459-460.

Tarn, W.Y. and Steitz, J.A. (1996) Highly diverged U4 and U6 small nuclear RNAs required for splicing rare AT-AC introns [see comments]. Science 273, 1824-1832.

ten Dijke, P., Ichijo, H., Franzen, P., Schulz, P., Saras, J., Toyoshima, H., Heldin, C.H. and Miyazono, K. (1993) Activin receptor-like kinases: a novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. Oncogene 8, 2879-2887.

ten Dijke, P., Franzen, P., Yamashita, H., Ichijo, H., Heldin, C.H. and Miyazono, K. (1994) Serine/threonine kinase receptors. [Review] [103 refs]. Progress. in Growth Factor. Research. 5, 55-72.

Terwilliger, J D and Ott J. (1994) Handbook of Human Genetic Linkage. Johns Hopkins University Press.

Van Cutsem, E., Rutgeerts, P., Geboes, K., Van Gompel, F. and Vantrappen, G. (1988) Estrogen-progesterone treatment of Osler-Weber-Rendu disease. Journal of Clinical Gastroenterology 10, 676-679.

Van Cutsem, E., Rutgeerts, P. and Vantrappen, G. (1990) Treatment of bleeding gastrointestinal vascular malformations with oestrogen-progesterone. Lancet 335, 953-955.

Van Cutsem, E. and Piessevaux, H. (1996) Pharmacologic therapy of arteriovenous malformations. [Review] [62 refs]. Gastrointestinal Endoscopy Clinics of North America 6, 819-832.

Vase, I. and Vase, P. (1979) Ocular lesions in hereditary haemorrhagic telangiectasia. Acta Ophthalmologica 57, 1084-1090.

Vase, P. (1981) Estrogen treatment of hereditary hemorrhagic telangiectasia. A double-blind controlled clinical trial. Acta Medica. Scandinavica. 209, 393-396.

Vase, P., Holm, M. and Arendrup, H. (1985) Pulmonary arteriovenous fistulas in hereditary hemorrhagic telangiectasia. Acta Medica. Scandinavica. 218, 105-109.

Vase, P. and Grove, O. (1986) Gastrointestinal lesions in hereditary hemorrhagic telangiectasia. Gastroenterology 91, 1079-1083.

Vase, P. and Lorentzen, M. (1983) Histological findings following oestrogen treatment of hereditary haemorrhagic telangiectasia. A controlled double-blind investigation. Journal. of. Laryngology. &. Otology. 97, 427-429.

Weber, F P. (1907) Multiple hereditary developmental angiomata (telangiectases) of the skin and mucous membranes associated with recurring haemorrhages. Lancet 2,160-162.

Ventura, F., Liu, F., Doody, J. and Massague, J. (1996) Interaction of transforming growth factor-beta receptor I with farnesyl-protein transferase-alpha in yeast and mammalian cells. Journal. of. Biological. Chemistry. 271, 13931-13934.

Vincent, P., Plauchu, H., Hazan, J., Faure, S., Weissenbach, J. and Godet, J. (1995) A third locus for hereditary haemorrhagic telangiectasia maps to chromosome 12q. Human. Molecular. Genetics 4, 945-949.

Wang, T., Donahoe, P.K. and Zervos, A.S. (1994) Specific interaction of type I receptors of the TGF-beta family with the immunophilin FKBP-12. Science 265, 674-676.

White, R.I., Jr. (1992) Pulmonary arteriovenous malformations: how do we diagnose them and why is it important to do so? [editorial; comment]. Radiology 182, 633-635.

White, R.I., Jr. (1996) Pulmonary arteriovenous malformations and hereditary hemorrhagic telangiectasia: embolotherapy using balloons and coils [letter; comment]. Archives. of. Internal. Medicine 156, 2627-2628.

White, R.I., Jr., Pollak, J.S. and Wirth, J.A. (1996) Pulmonary arteriovenous malformations: diagnosis and transcatheter embolotherapy. [Review] [56 refs]. Journal. of. Vascular. &. Interventional. Radiology 7, 787-804.

White, T, Arnheim N and Erlich H. (1989) The polymerase chain reaction. Trends in Genetics 5, 185-189.

Yamamoto, H., Ueno, H., Ooshima, A. and Takeshita, A. (1996) Adenovirus-mediated transfer of a truncated transforming growth factor-beta (TGF-beta) type II receptor completely and specifically abolishes diverse signaling by TGF-beta in vascular wall cells in primary culture. Journal of Biological Chemistry 271, 16253-16259.

Yamashita, H., Ichijo, H., Grimsby, S., Moren, A., ten Dijke, P. and Miyazono, K. (1994a) Endoglin forms a heteromeric complex with the signaling receptors for transforming growth factor-beta. Journal. of. Biological. Chemistry. 269, 1995-2001.

Yamashita, H., ten Dijke, P., Franzen, P., Miyazono, K. and Heldin, C.H. (1994b) Formation of hetero-oligomeric complexes of type I and type II receptors for transforming growth factor-beta. Journal. of. Biological. Chemistry. 269, 20172-20178.

Zhang, Y., Feng, X., We, R. and Derynck, R. (1996) Receptor-associated Mad homologues synergize as effectors of the TGF-beta response. Nature 383, 168-172.

□

Appendix A. Genbank submission of sequence generated during this work.

```
LOCUS       HSALK01      460 bp    DNA             PRI       01-JUL-1997
DEFINITION  Human activin receptor like kinase 1 (ALK-1) gene, exon 2.
ACCESSION   U77707
NID         g2228554
KEYWORDS    .
            1 of 7
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 460)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     The activin receptor like kinase 1 gene - genomic structure of the
            coding region
  JOURNAL   Am. J. Hum. Genet. (1997) In press
REFERENCE   2  (bases 1 to 460)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1996) Department of Genetics, Duke University,
            CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..460
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            254..319
                     /gene="ALK-1"
                     /number=2
BASE COUNT       99 a    132 c    135 g     94 t
ORIGIN
        1 tttccccaag ctctcagtca cttaaacatt gctctccacc cttcacctct aacaggatgg
       61 tttccatggg gaagtgaacc aggacttccc ctgcaggccc cgccccaaag ccaggcggca
      121 gggagtaggg aggcggcctc cctgcctccc ctccaaaaaa aactctgtga tttcctctgg
      181 gcaggaggga gccacggcca gcggctgtca cacttcatgg ctcttactcc acctctcttg
      241 ctcctctctg cagggaccat gaccttgggc tcccccagga aaggccttct gatgctgctg
      301 atggccttgg tgacccaggg tgagtactgg gggagcagtt aggaaacagg aacctggata
      361 cagaaagggc tatctgggcc cagatcagct ctgcctgggg ctgaacttga gaagctgggg
      421 agaatgtagg agcttgactg gagagtggag gacagtgagg
//
```

```
LOCUS       HSALK02     1371 bp    DNA             PRI      01-JUL-1997
DEFINITION  Human activin receptor like kinase 1 (ALK-1) gene, exons 3, 4 and
            5.
ACCESSION   U77708
NID         g2228555
KEYWORDS    .
            2 of 7
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 1371)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     The activin receptor like kinase 1 gene - genomic structure of the
            coding region
  JOURNAL   Am. J. Hum. Genet. (1997) In press
REFERENCE   2  (bases 1 to 1371)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1996) Department of Genetics, Duke University,
            CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..1371
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            160..411
                     /gene="ALK-1"
                     /note="encodes the extracellular domain"
                     /number=3
     exon            619..830
                     /gene="ALK-1"
                     /note="encodes the extracellular domain and transmembrane
                     domain"
                     /number=4
     exon            1034..1133
                     /gene="ALK-1"
                     /note="encodes GS domain"
                     /number=5
BASE COUNT     246 a    353 c    516 g    256 t
ORIGIN
        1 gttctaggga aggatgactg aggatgaaag taagagacca aagcttcaag gtgtttgtct
       61 gaggggtcag acgagaggga cagtaggaca gaaatgggtg tcgggctcag cctgggggag
      121 ctgggaccac agtggctgag cttccggtgt gtcttccagg agaccctgtg aagccgtctc
      181 ggggcccgct ggtgacctgc acgtgtgaga gcccacattg caaggggcct acctgccggg
      241 gggcctggtg cacagtagtg ctggtgcggg aggaggggag gcacccccag gaacatcggg
      301 gctgcgggaa cttgcacagg gagctctgca gggggcgccc caccgagttc gtcaaccact
      361 actgctgcga cagccacctc tgcaaccaca acgtgtccct ggtgctggag ggtacgtcca
      421 gctgccctag cactccctcc ccatcttctt ggcccctgcc tcccttccct cctttcctct
      481 catgctctgg ccaataaagg ggctgggggc gggggagcgg gtgggcagga ctctgggatc
      541 taactggcag agtggtctgg cccgaggtgg ggggagctga cctagtggaa gctgagcctc
      601 agtgtccccc tccctcagcc acccaacctc cttcggagca gccgggaaca gatggccagc
      661 tggccctgat cctgggcccc gtgctggcct tgctggccct ggtggccctg ggtgtcctgg
      721 gcctgtggca tgtccgacgg aggcaggaga agcagcgtgg cctgcacagc gagctgggag
      781 agtccagtct catcctgaaa gcatctgagc agggcgacag catgttgggg gtatgggcct
      841 ggggacctgg gacacagggt gtaggagggg cagataggaa ctgcagaatc agaggggtca
```

901 cccagagatt agagccggtg gggagctggg cgagtgagga gcttgcagtg acccagcagg
961 tcccaggtcg aggatagaga agggggctgt ggctggttgt ggcagcctct cagtggcctc
1021 tccgtacccc caggacctcc tggacagtga ctgcaccaca gggagtggct cagggctccc
1081 cttcctggtg cagaggacag tggcacggca ggttgccttg gtggagtgtg tgggtgagca
1141 gtgggtgagc ccggtggatg aggaccaagg gctctcatga gctggagggg tgagggagtt
1201 tttggctact ggaatcacag gcggtgcagg cctgggtcag ggagtgggct ggagacgggc
1261 cagggctagg ttcttctttc tgcaggaccg gggtggaacg agaggcagct gggggtggcc
1321 tgccactggg tttgggtctg gattaagtta aacctaaggg tctggggttc t
//

```
LOCUS       HSALK03      317 bp    DNA             PRI      01-JUL-1997
DEFINITION  Human activin receptor like kinase 1 (ALK-1) gene, exon 6.
ACCESSION   U77709
NID         g2228556
KEYWORDS    .
            3 of 7
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 317)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     The activin receptor like kinase 1 gene - genomic structure of the
            coding region
  JOURNAL   Am. J. Hum. Genet. (1997) In press
REFERENCE   2  (bases 1 to 317)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1996) Department of Genetics, Duke University,
            CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..317
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            89..235
                     /gene="ALK-1"
                     /note="encodes an ATP binding site and a kinase domain;
                     the 3' splice junction of exon 6 has a non-consensus
                     sequence AGgcaag, this has been reported previously in
                     other genes"
                     /number=6
BASE COUNT       67 a    84 c   101 g    64 t     1 others
ORIGIN
        1 agggaggcag cgcagcatca agatgggggg ctcttccagg gctctgtgtg cccagtgtgt
       61 aaccctcacc ttcccctctg gccatcagga aaaggccgct atggcgaagt gtggcggggc
      121 ttgtggcacg gtgagagtgt ggccgtcaag atcttctcct cgagggatga acagtcctgg
      181 ttccgggaga ctgagatcta taacacagtg ttgctcagac acgacaacat cctaggcaag
      241 gggagaggcc agctgtgcca ggcctggggc tttgccnctg cactcagggc tcaagtttgc
      301 agacctccag acattaa
//
```

LOCUS HSALK04 441 bp DNA PRI 01-JUL-1997
DEFINITION Human activin receptor like kinase 1 (ALK-1) gene, exon 7.
ACCESSION U77710
NID g2228557
KEYWORDS .
4 of 7
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE The activin receptor like kinase 1 gene - genomic structure of the
coding region
JOURNAL Am. J. Hum. Genet. (1997) In press
REFERENCE 2 (bases 1 to 441)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1996) Department of Genetics, Duke University,
CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES Location/Qualifiers
source 1..441
/organism="Homo sapiens"
/db_xref="taxon:9606"
exon 108..383
/gene="ALK-1"
/note="encodes a kinase domain"
/number=7
BASE COUNT 81 a 166 c 115 g 79 t
ORIGIN
1 cagtgaccca gtccattccc tctcccccaa ccccaccctg accctgacga ctccagcctc
61 ccttagcccc agccccttgc tgagtcaccc aacctttctg cacacaggct tcatcgcctc
121 agacatgacc tcccgcaact cgagcacgca gctgtggctc atcacgcact accacgagca
181 cggctccctc tacgactttc tgcagagaca gacgctggag ccccatctgg ctctgaggct
241 agctgtgtcc gcggcatgcg gcctggcgca cctgcacgtg gagatcttcg gtacacaggg
301 caaaccagcc attgcccacc gcgacttcaa gagccgcaat gtgctggtca agagcaacct
361 gcagtgttgc atcgccgacc tgggtgagcc gggcggggca ggggcgcgcc cttcacaggt
421 gggcggagct tgtgcgctct c
//

```
LOCUS       HSALK05      522 bp    DNA             PRI       01-JUL-1997
DEFINITION  Human activin receptor like kinase 1 (ALK-1) gene, exon 8.
ACCESSION   U77711
NID         g2228558
KEYWORDS    .
            5 of 7
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
            Homo.
REFERENCE   1  (bases 1 to 522)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     The activin receptor like kinase 1 gene - genomic structure of the
            coding region
  JOURNAL   Am. J. Hum. Genet. (1997) In press
REFERENCE   2  (bases 1 to 522)
  AUTHORS   Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-1996) Department of Genetics, Duke University,
            CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1..522
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            160..357
                     /gene="ALK-1"
                     /note="encodes a kinase domain"
                     /number=8
BASE COUNT       93 a    168 c    151 g    109 t      1 others
ORIGIN
        1 atggttctct ctgtggccac tgccttccag cccatctccg tgcacgtctc catctgcctt
       61 cccctctctg tcccactgtt tctctcagtc cccaccttgc ctgccccctg gatcccaggt
      121 ttgggagagg ggcaggagtg acaggcctca cccccacagg cctggctgtg atgcactcac
      181 agggcagcga ttacctggac atcggcaaca acccgagagt gggcaccaag cggtacatgg
      241 cacccgaggt gctggacgag cagatccgca cggactgctt tgagtcctac aagtggactg
      301 acatctgggc ctttggcctg gtgctgtggg agattgcccg ccggaccatc gtgaatggtg
      361 agggcccacc ctacacaggg tagggaaagg ggaatcagcc tgtggagcca ggggcttcca
      421 gccatggcca gtgcccatgg cctggnaggt ttgcagtcag acctcctggc acccccttcca
      481 tgctgcccac cagctggttc agctgagtga ccttttaagg ta
//
```

LOCUS HSALK06 413 bp DNA PRI 01-JUL-1997
DEFINITION Human activin receptor like kinase 1 (ALK-1) gene, exon 9.
ACCESSION U77712
NID g2228559
KEYWORDS .
6 of 7
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE The activin receptor like kinase 1 gene - genomic structure of the
coding region
JOURNAL Am. J. Hum. Genet. (1997) In press
REFERENCE 2 (bases 1 to 413)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1996) Department of Genetics, Duke University,
CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES Location/Qualifiers
source 1..413
/organism="Homo sapiens"
/db_xref="taxon:9606"
exon 147..277
/gene="ALK-1"
/note="encodes a kinase domain"
/number=9
BASE COUNT 73 a 108 c 123 g 109 t
ORIGIN
1 ttgcattata ctgtccctct caggggtagc gtgtccaggc cactggtttc tggcccttgg
61 atagagggta gaaaaggctc tcctctgggt ggtattgggc ctccttagag tcccaagtga
121 ttgtcctgtc cattctccat ttccaggcat cgtggaggac tatagaccac ccttctatga
181 tgtggtgccc aatgacccca gctttgagga catgaagaag gtggtgtgtg tggatcagca
241 gacccccacc atccctaacc ggctggctgc agacccggtg aggcctctgc tgggactagg
301 atggcgtggg gtggtggctc acggctggga tttctgggcc cgggaacttg tgtctgaggc
361 ctctgcttca tctcatagat actgagtgtc ctggttaggc acctctctag tac
//

LOCUS HSALK07 264 bp DNA PRI 01-JUL-1997
DEFINITION Human activin receptor like kinase 1 (ALK-1) gene, exon 10 and
complete cds.
ACCESSION U77713
NID g2228560
KEYWORDS .
7 of 7
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae;
Homo.
REFERENCE 1 (bases 1 to 264)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE The activin receptor like kinase 1 gene - genomic structure of the
coding region
JOURNAL Am. J. Hum. Genet. (1997) In press
REFERENCE 2 (bases 1 to 264)
AUTHORS Berg,J.N., Gallione,C., Stenzel,T. and Marchuk,D.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1996) Department of Genetics, Duke University,
CARL Building, Research Drive, Durham, NC 27710, USA
FEATURES Location/Qualifiers
source 1..264
/organism="Homo sapiens"
/db_xref="taxon:9606"
mRNA join(U77707:<1..319,U77708:160..411,U77708:619..830,
U77708:1034..1133,U77709:89..235,U77710:108..383,
U77711:160..357,U77712:147..277,98..>264)
/product="activin receptor like kinase 1"
gene join(U77707:254..460,U77708:1..1371,U77709:1..317,
U77710:1..441,U77711:1..522,U77712:1..413,1..264)
/gene="ALK-1"
CDS join(U77707:259..319,U77708:160..411,U77708:619..830,
U77708:1034..1133,U77709:89..235,U77710:108..383,
U77711:160..357,U77712:147..277,98..232)
/gene="ALK-1"
/codon_start=1
/product="activin receptor like kinase 1"
/db_xref="PID:g2228562"

/translation="MTLGSPRKGLLMLLMALVTQGDPVKPSRGPLVTCTCESPHCKGP

TCRGAWCTVVLVREEGRHPQEHRGCGNLHRELCRGRPTEFVNHYCCDSHLCNHNVSLV

LEATQPPSEQPGTDGQLALILGPVLALLALVALGVLGLWHVRRRQEKQRGLHSELGES

SLILKASEQGDSMLGDLLDSDCTTGSGSGLPFLVQRTVARQVALVECVGKGRYGEVWR

GLWHGESVAVKIFSSRDEQSWFRETEIYNTVLLRHDNILGFIASDMTSRNSSTQLWLI

THYHEHGSLYDFLQRQTLEPHLALRLAVSAACGLAHLHVEIFGTQGKPAIAHRDFKSR

NVLVKSNLQCCIADLGLAVMHSQGSDYLDIGNNPRVGTKRYMAPEVLDEQIRTDCFES

YKWTDIWAFGLVLWEIARRTIVNGIVEDYRPPFYDVVPNDPSFEDMKKVVCVDQQTPT

IPNRLAADPVLSGLAQMMRECWYPNPSARLTALRIKKTLQKISNSPEKPKVIQ"

exon 98..>264
/gene="ALK-1"
/note="encodes a kinase domain"
/number=10
BASE COUNT 55 a 104 c 45 g 60 t
ORIGIN
1 tgacccacct ccctctgcat ctctctcccg acccccctcct cttctctgca tctctctctc
61 tgcctcctct cctctgcacc tctctcccaa cccccaggtc ctctcaggcc tagctcagat
121 gatgcgggag tgctggtacc caaacccctc tgcccgactc accgcgctgc ggatcaagaa
181 gacactacaa aaaattagca acagtccaga gaagcctaaa gtgattcaat agcccaggag
241 cacctgattc ctttctgcct gcag
//

Appendix B.

Clinical data from families 5,6,8 and 9 collected during the study.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	Gastro-intestinal Haemorrhage	Other	Diagnosis of HHT Yes/No/Uncertain
2090	86	Yes, weekly	None	Lips, ears, tongue		No	No	Yes
2091	55	No		No		No	No	No
2092	67	yes, Daily	Laser Septodermoplasty	Lips, ears, tongue, fingers	No	No	No	Yes
2094	81	Very rarely	None	None		No	No	No
2096	27	Occasional	None	No		1 episode of colonic bleeding	No	Uncertain
2097	21	Occasional	None	No	No	No	No	Uncertain
2098	26	As Child, not since	None	No	No	No	No	No
2099	49	No	None	No	No	No	No	No
2100	35	Occasional in pregnancy	None	No	No	No	No	No
2101	38	No	None	No	No	No	No	No
2102	75	Yes, frequent	Septodermoplasty (twice)	Face, hands, lips, tongue	No	No	No	Yes
2103	69	Yes, Frequent	Cautery (unsuccessful)	Face, Lips, Ears	No	No	No	Yes
2104	41	As a child, not since	None	No	No	No	No	No

Clinical details of family 9. The individual numbers are those given on the family tree given in figure 3.1. All family members were screened for PAVMs by pulse oximetry, which showed no evidence of a PAVM in any family member. The site of telangiectases is given, as are the patients description of frequency and severity of nosebleeds. Diagnosis of HHT is based on the criteria given in the Materials and Methods section. Continued overleaf.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	Gastro-intestinal Haemorrhage	Other	Diagnosis of HHT Yes/No/Uncertain
2105	52	Daily	None	Single lesion on nose	No	No	No	Yes
2106	26	No	None	No	No	No	No	No
2107	23	As a child, rarely since.	None	No	No	No	No	No
2109	44	Occasional	None	A few spider naevi only	No	No	No	No
2111	81	Previously severe, now resolved	None	Lips, tongue, ears, fingers.	No	No	No	Yes
2113	31	Frequent	None	Lips and ears	No	No	Brain Haemorrhage	Yes
2114	41	Frequent	None	In nose only	No	No	No	Yes
2115	34	Daily	None	Tongue and Ears	No	No	No	Yes
2116	77	Daily, moderate	Septodermoplasty Cautery	Fingers and Lips	No	No	No	Yes
2118	45	2 to 3 moderate bleeds per day	None	Single lesion on ear	No	Episodes rectal bleeding	No	Yes
2119	53	Occasionally, worse as youth	None	Suspect lesions on lower lip	No	No	No	Uncertain
2120	74	One to two per month.	None	Suspect lesions on fingers + lips	No	No	No	Yes
2121	44	None	None	Single suspect lesion on lip	No	No	No	no

Clinical details of family 9

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	Gastro-intestinal Haemorrhage	Other	Diagnosis of HHT Yes/No/Uncertain
2122	50	Occasional (less than monthly)	None	No	No	No	No	No
2123	62	Occasional	None	No	No	No	No	No
2125	73	None	None	No	No	No	No	No
2129	21	Very Occasional	None	No	No	No	No	No
2318	58	Daily,mild	Cautery	Lips, Tongue, Fingers	Yes (visual aura)	Yes (No details)	No	Yes
2336	32	Weekly, mild	None	No	No	No	No	Yes
2337	86	Weekly, Moderate	None	Lips, tongue, ears, fingers	No	No	No	Yes
2340	65	Less than monthly, mild	None	No	No	No	No	No
2342	63	At least twice daily	Septodermoplasty	Lips, tongue, ears, fingers	No	No	No	Yes
2343	52	During pregnancy	None	Lips, ears (sparse)	Yes	No	Unexplained haematuria	Yes
2347	62	Mild, daily	None	Lips, tongue, ears	No	No	No	Yes
2349	31	None	None	No	Occasional headaches	No	No	No
2350	55	None	None	Few on lips	No	unexplained anaemia	No	Yes
2352	55	Weekly, moderate	None	Multiple on lips and ears	No	No	No	Yes

Clinical details of family 9 (continued). Individuals 2318 to 2352 come from a second branch of the same family as discussed in the text.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	PAVMs	Other	Diagnosis of HHT Yes/No/Uncertain
2171	60	Daily, lasting 5 to 15 minutes each	Cautery (not effective)	Lips, tongue, ears, fingers	No	No	Elevated liver function tests	Yes
2255	51	Occasional, very minor	None	No	No	No	No	No
2256	32	None	None	No	No	No	No	No
2257	54	As youth, none since	None	No	No	No	No	No
2259	34	Approximately yearly	None	Single suspect lesion on lip	Yes	No	No	Uncertain
2260	15	Less than Monthly	None	No	Yes	No	No	Uncertain
2261	49	None	None	No	No	No	No	No
2262	56	Daily, severe	Blood transfusion	All over lips, face, hands.	No	Probable (Low Oxygen Saturation)	Recurrent G-I blood loss	Yes
2263	33	None	None	No	No	No	No	No
2265	39	None	None	No	No	No	No	No
2266	27	None	None	No	No	No	No	No
2267	41	Weekly	None	3 small lesions on lips	No	Single apical PAVM.	No	Yes
2268	38	None	None	No	No	No	No	No
2269	7	Weekly	None	No	No	No	No	No

Clinical details of family 8. Individual numbers are given on the family tree in figure 3.1. All family members were screened by pulse oximetry

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	PAVMs	Other	Diagnosis of HHT Yes/No/Uncertain
2270	12	None	None	No	No	No	No	No
2271	34	Weekly	None	Sparsely on tongue and lips	No	Yes, haemoptysis in pregnancy	No	Yes
2273	44	Weekly	None	Tongue and lips	No	Probable (low oxygen saturation)	No	Yes
2274	18	Weekly	None	No	No	No	No	Yes
2275	45	Weekly, severe	None	3 on lip, small venous lesion on arm	Yes	No	No	Yes
2276	32	None	None	No	No	No	No	No
2277	40	Weekly	None	Small lesions on tip of tongue	No	No	No	Yes
2278	13	Weekly	None	No	No	No	No	Yes
2280	37	Occasional	None	No	No	No	No	No
2282	41	None	None	Marked spider naevi no telangiectasis	No	No	No	No
2283	47	None	None	No	No	No	No	No
2284	35	None	None	No	No	No	No	No
2285	63	None	None	No	No	No	No	No

Clinical details of family 8 (continued). Individual numbers are given on the family tree in figure 3.1.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	Gastro-intestinal Haemorrhage	Other	Diagnosis of HHT Yes/No/Uncertain
9196	53	Monthly for past 15 years.	None	Single suspect lesion on palate	No	No	No	Uncertain
9199	51	None	None	No	No	No	No	No
9200		Daily,More than 15 minutes duration	None	Lips, tongue, fingers	No	Yes (5 gastric lesions seen on endoscopy)	No	Yes
9201	55	Daily, 5 to 15 minutes	None	Lips, tongue, fingers, palate, nailbeds	Yes	Gastric telangiectases on endoscopy	No	Yes
9203	27	Less than monthly	None	Single probable lesion on nailbed	Yes	No	No	Yes
9204	50	Weekly, lasting 5 minutes each.	None	4 lesions on lips	Simple migraine	No	No	Yes
9208	71	Daily, prolonged	Cautery Dermoplasty	Lips, tongue, palate, fingers, nailbeds	Yes	No	No	Yes
9209	26	Less than monthly	None	No	Yes	No	No	Uncertain
9210	48	Only as a child	None	No	No	No	No	No
9211	39	Monthly, brief episodes	None	No	No	No	No	Yes
9212	30	Monthly, fortnightly when pregnant	None	Hands and right side of face.	No	No	No	Yes

Clinical details of family 5. Individual numbers are given on the family tree in figure 3.1. All family members were screened by pulse oximetry, with CXR and lung scan of individuals 9214, 9215 and 9196 who had an oxygen saturation drop between supine and erect position. No PAVM was found. No evidence of a PAVM was present in any other family member. Individual 9207 was an affected individual from the same region, thought to be a relative of the family. Table is continued overleaf.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	Gastro-intestinal Haemorrhage	Other	Diagnosis of HHT Yes/No/Uncertain
9213	30	Weekly, 5 to 15 minutes duration	None	No	Yes	No	No	Yes
9214	71	More than daily, last more than 15 minutes.	Repeated packing	Lips, tongue, palate, fingers, nailbeds	Yes	No	Hepatic lesion	Yes
9215	52	More than daily, onset at age of 12	Cautery (ineffective)	lips, tongue, nailbeds, fingers	Yes	No	No	Yes
9216	24	Daily, more than 5 minutes	cautery (ineffective)	Single lesion on lip	Yes	No	No	Yes
9217	49	None	None	No	No	No	No	No
9207	77	Frequent, less than 5 minutes duration		Lips, tongue and fingers	No	Haematemesis and melena (?origin)	No	Yes

Clinical data from family 5 continued.

Individual	Age	Nosebleeds Frequency	Nosebleeds Treatment	Telangiectasis	Classical Migraine	PAVMs	Other	Diagnosis of HHT Yes/No/Uncertain
70	55	Several times a day	Cautery	Tongue and lips	Daily	Lesions causing haemoptysis	No	Yes
72	32	None	None	No	No	No	No	No
73	31	Weekly since childhood	None	Tongue and lower lip	Yes	No	No	Yes
74	51	Daily, prolonged	None	Tongue and lips	Weekly	Suspect lesion on chest X-ray	CAVM treated with radiosurgery	Yes
75	31	Less than monthly, lasting 5 minutes	None	On Tongue	No	No	No	Yes
76	30	None	None	No	No	No	Haematuria once. Oesophagitis	No
77	26	Occasionally	None	Tongue lesions	No	No	No	Yes

Clinical data from family 6. The sister of individuals 70 and 74 died of a cerebral haemorrhage caused by a cerebral arteriovenous malformation

Appendix C. Questionnaire used to collect patient data during family visits.

Full Name: .....

Date of Birth:

Address: **telephone**

General Practitioner .....

Address:

---

1. Approximately how old were you when you first had problems with HHT ?

2. Please tick the first sign

Nosebleeds [ ]
Red spots [ ]
Other [ ]

please specify .....

3. Do you get nose bleeds **Yes / No**

a.) If so, how often do they happen:

- once or more daily [ ]
- once or more weekly [ ]
- once or more monthly [ ]
- less than monthly [ ]

b.) Is the bleeding worst from

- Right Nostril [ ]
- Left Nostril [ ]
- No Difference [ ]

c.) How long do they last

- less than 5 minutes [ ]
- 5 to 15 minutes [ ]
- Longer than 15 minutes [ ]

d.) Have you had any treatment for them **Yes / No**

e.) If so, what treatment ?

- cautery [ ]
- Laser Treatment [ ]
- Operation [ ]
- Other [ ]

if other please specify .....

4. Do you get indigestion ? **Yes / No**

a.) If so, how often ?

Daily [ ]
Weekly [ ]
Less often [ ]

b.) Do you take any treatment for it ? **Yes / No**

If so, what do you take ?

c.} Have you ever had an ulcer ? **Yes / No**

5. Have you ever had any bleeding into your digestive system ? Yes/No

a.) If so, have you had any treatment for it ?

b.)What treatment have you had ?

Laser treatment [ ]
Operation [ ]
Other [ ]

If other, please specify.

6. Have you had any of the following tests ?

Endoscopy Yes / No
Barium meal Yes / No

If so, was anything found ?

Ulcer [ ]
Dilated blood vessels [ ]
Other [ ]

7.Are you on any other treatment for digestive problems ? Yes/No

If so, please specify.....

8.Have you ever had any chest problems ? **Yes / No**

a.)Please tick any in the list below :

"Abnormal Blood Vessels" (arteriovenous malformations) [ ]
Blood in your sputum [ ]
breathlessness on exercise [ ]
Other [ ]

please specify.....

b.)Have you ever had an Angiogram ? **Yes/No**

c.)Have you ever had embolisation treatment ? **Yes/No**

If so, what year did you first have it ?

How often have you had it since ?

9. Have you ever had a chest X-ray ? **Yes/No**

Did it show anything ? .....

10 Do you get headaches ? **Yes/No**

a.) If so, how often do they happen

Daily [ ]
Weekly [ ]
Monthly [ ]
Rarely. [ ]

b.) Do you ever get visual problems before a headache ? **Yes/No**

c.) Do you have any treatment for your headaches?

Are there any other problems you think are important ?

Part II

Epistaxis: Frequency
Worse side:
Severity:
Treatment:

Gastrointestinal: Indigestion
Ulcers
Haematemesis
Melena
Treatment

PAVM
CXR
Treatment

Cardiac Failure:

Migraine: 2 attacks aura symptom 4-60mins headache follows odc fulfilled y/n

CAVM

Stroke

Aneurysm

Miscarriages

On Examination

Telangiectases: lips tongue palate nailbeds fingers

gums conjunctivae Elsewhere:

Haemangiomata:

Oxygen Saturation: supine standing

Any other findings:

Diagnosis of HHT: Yes / No / Uncertain
-blood taken for DNA blood sent for transformation

-photos taken :frame no + site.

Consent Form.

I consent to give blood for research into Hereditary Haemorrhagic Telangiectasia.

I am happy for my records to be stored on a computer database.

I am happy for a photograph to be taken.

I am happy for the photograph to be used in a scientific publication (steps will be taken to make the picture anonymous.)

Signature

Name . Date.

Appendix D. Publications in which work from this thesis is included.

J Med Genet 1996;33:256–257

Clinical heterogeneity in hereditary haemorrhagic telangiectasia: are pulmonary arteriovenous malformations more common in families linked to endoglin?

Jonathan N Berg, Alan E Guttmacher, Douglas A Marchuk, Mary E M Porteous

Abstract

Pulmonary arteriovenous malformations (PAVMs) occur in up to 27% of patients with hereditary haemorrhagic telangiectasia (HHT) and are associated with a rate of paradoxical cerebral embolism at presentation of up to 36%. At least two different loci have been shown for HHT. Mutations in endoglin have been found in some families and the locus designated ORW1. In other families this locus has been excluded. In this paper we confirm that in families linked to ORW1 there is a prevalence of PAVMs among affected members of 29·2%, compared to a prevalence of 2·9% in families in which this locus has been excluded ($\chi^2 = 19·2$, $p < 0·001$). This information can be used to decide how to screen HHT patients for PAVMs.

(*J Med Genet* 1996;33:256–257)

Key words: hereditary haemorrhagic telangiectasia; pulmonary arteriovenous malformation; endoglin.

Human Genetics Unit, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, UK
J N Berg
M E M Porteous

Department of Pediatrics, University of Vermont College of Medicine, Burlington, VT 05401, USA
A E Guttmacher

Department of Genetics, Duke University Medical Centre, Durham, NC 27710, USA
D A Marchuk

Correspondence to:
Dr Berg.

Received 23 May 1995
Revised version accepted for publication 16 October 1995

Hereditary haemorrhagic telangiectasia (HHT), also known as Osler-Rendu-Weber syndrome, is an autosomal dominant vascular dysplasia characterised by mucocutaneous telangiectasis (fig 1) and severe recurrent epistaxes. Pulmonary arteriovenous malformations (PAVMs) occur in between 4·8% and 27%[12] of patients. PAVMs may be asymptomatic, but can often cause hypoxaemia and pulmonary haemorrhage. PAVMs also lead to high incidence of cerebral abscesses, and cerebral thromboembolic complications occur in up to 36% of patients at presentation.[3] Most authors recommend screening for asymptomatic PAVMs in HHT patients,[3–5] with occlusion by embolisation of all lesions with a feeding vessel greater than 3 mm.

Two loci for HHT have been identified. The ORW1 locus, mapping to chromosome 9q34, is endoglin, a TGF-β binding protein expressed on endothelial cells.[6–9] Recently, a second locus on 12q has been reported.[10 11] Several authors have noted a subjectively higher frequency of PAVMs in families showing linkage to ORW1.[8 12 13]

In this paper we aim to address the size of this difference, using data from three new families (fig 2) and 16 families previously published.[7 10 11]

Diagnosis of HHT in the three new families was made by the presence of two of: recurrent epistaxis, mucocutaneous telangiectasis, and an affected first degree relative. Pulse oximetry was performed according to the method described by Hughes[4] to detect PAVMs. Three subjects from family 5 had a chest *x* ray and lung perfusion scans. Clinical details of the thirteen families assessed by our group were all collected before linkage data were available.

For the newly assessed families, microsatellite markers were run using accepted techniques and analysis of results was performed using MLINK or LINKMAP.

Of the 16 families included, eight families are not linked to endoglin on chromosome 9

Families used in this study: source of data and incidence of PAVMs within each family

Family source and reference	*Endoglin (ORW1) linkage*	*Assessed by pulse oximetry*	*No of affected family members assessed*	*No of PAVMs in those assessed*
Family 1[12]	No	Yes	6	0
Family 2[12]	No	Yes	8	0
Family 3[12]	No	No	10	0
Family 4[12]	No	No	9	1
Family 5*	No	Yes	13	0
Family 17[8]	No	No	6	0
Family 33[8]	No	No	8	0
Family T[7]	No	Yes	9	1
		Total	69	2 (2·9%)
Family 01[8]	Yes†	No	45	13
Family 02[8]	Yes†	No	20	8
Family 32[8]	Yes†	No	11	3
Family 56[8]	Yes‡	No	6	2
Family 6*	Yes‡	Yes	5	2
Family 7*	Yes‡	Yes	7	3
Family F[7]	Yes‡	Yes	9	3
Family A[7]	Yes†	Yes	17	1
		Total	120	35 (29·2%)

In non-ORW1 linked families linkage was excluded across the candidate interval on multipoint analysis, or by a lod of less than −2 at $\theta = 0·02$ with D9S61.
* Family previously unpublished (Porteous and Berg).
† ORW1 linkage was shown by a lod score of greater than 3 to a marker close to the candidate region.
‡ Lod score of greater than 1 at $\theta = 0$ with D9S61, D9S63, or D9S65, close to the maximum lod obtainable with the family.

Figure 1 Classical telangiectases on the lips of an HHT patient.

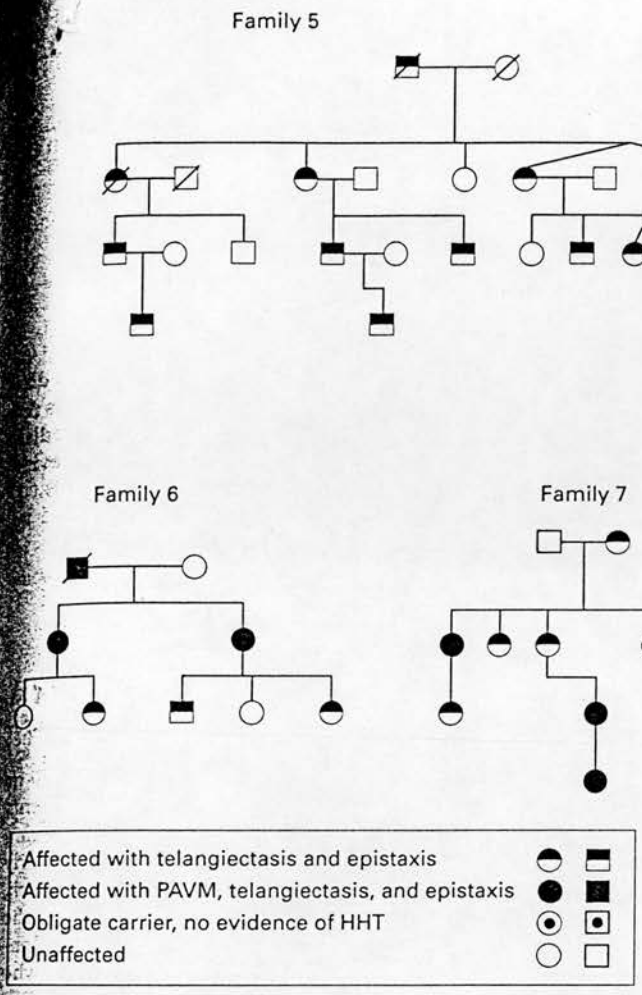


Figure 2 Pedigrees of the three families not previously published.

and eight families show evidence of linkage. The majority of people in this group (93/120) came from families large enough to achieve a lod score of greater than 3. The data are summarised in the table.

In the families linked to endoglin, 35/120 (29·2%) had pulmonary arteriovenous malformations. In the unlinked families, 2/69 (2·9%) had PAVMs. The difference between these two groups is highly significant ($\chi^2 = 19{\cdot}2$, 1 df, $p<0{\cdot}001$).

In the endoglin linked group many of the patients had symptomatic PAVMs. There may be several small asymptomatic PAVMs in the non-endoglin population, even though 36/69 were screened by pulse oximetry. This is an unlikely source of bias as more careful screening of the endoglin population would also be expected to turn up a greater number of asymptomatic PAVMs. All families were assessed "blind" before linkage data were available.

The observed difference in PAVM frequency is a biologically plausible finding. Endoglin may well have a role in the development of pulmonary vasculature which is not shared by the other genes in which mutations can cause HHT.

As all those with HHT are at some risk of having a PAVM, it is important to offer screening to all patients, at least by the method of Hughes[4] using supine and erect pulse oximetry and chest *x* ray.

With members of endoglin linked families having a 29·2% incidence of PAVMs, it is particularly important to target screening on (1) patients from families with HHT linked to endoglin on 9q34, (2) patients who have a family history of PAVMs. For these patients, we feel that formal arterial oxygen measurement and measurement of right to left pulmonary shunt fraction using the 100% oxygen method[5] may be justified.

Firstly we wish to thank all the patients whose participation made this study possible. We also wish to thank Professor J M B Hughes for assessing members of family 5 for PAVMs and Bernadette Farren for helping to collect samples. Particular thanks are due to Catherine Cookson, who has funded much of this work.

1 Porteous MEM, Burn J, Proctor SJ. Hereditary haemorrhagic telangiectasia: a clinical analysis. *J Med Genet* 1992;29:527–30.
2 Plauchu H, de Chadarevian JP, Bideau A, *et al.* Age-related clinical profile of hereditary haemorrhagic telangiectasia in an epidemiologically recruited population. *Am J Med Genet* 1989;32:291–7.
3 White RI Jr, Lynch-Nyhan A, Terry P, *et al.* Pulmonary arteriovenous malformations: techniques and long term outcome of embolotherapy. *Radiology* 1988;169:663–9.
4 Hughes JMB. Intrapulmonary shunts: coils to transplantation. *J R C Physicians Lond* 1994;28:247–53.
5 Haitjema TJ, Overtoom TTh, Westermann CJJ, *et al.* Embolisation of pulmonary arteriovenous malformations: results and follow up in 32 patients. *Thorax* 1995;50:719–23.
6 McDonald MT, Papenberg KA, Ghosh S, *et al.* A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33–34. *Nature Genet* 1994;6:197–204.
7 Shovlin CL, Hughes JMB, Tuddenham EGD, *et al.* A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genet* 1994;6:205–9.
8 McAllister KA, Grogg KM, Johnson DW, *et al.* Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genet* 1994;8:345–51.
9 Heutink P, Haitjema T, Breedveld GJ, *et al.* Linkage of hereditary haemorrhagic telangiectasia to 9q34 and evidence for locus heterogeneity. *J Med Genet* 1994;31:933–6.
10 Vincent P, Plauchu H, Hazan J, *et al.* A third locus for hereditary haemorrhagic telangiectasia maps to 12q. *Hum Mol Genet* 1995;5:945–50.
11 Johnson DW, Berg JN, Gallione C, *et al.* A second locus for hereditary haemorrhagic telangiectasia maps to chromosome 12. *Genome Res* 1995;5:21–8.
12 Porteous MEM, Curtis A, Williams O, *et al.* Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *J Med Genet* 1994;31:925–6.
13 McAllister KA, Lennon F, Bowles-Bieseckker B, *et al.* Genetic heterogeneity in hereditary haemorrhagic telangiectasia: possible correlation with clinical phenotype. *J Med Genet* 1994;31:927–32.

Human Molecular Genetics, 1995, Vol. 4, No. 10 1983–1985

Six novel mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type 1 suggest a dominant-negative effect of receptor function

Kimberly A.McAllister, Melanie A.Baldwin, Arun K.Thukkani, Carol J.Gallione, Jonathan N.Berg[1], Mary E.Porteous[1], Alan E.Guttmacher[2] and Douglas A.Marchuk*

Department of Genetics, Duke University Medical Center, Durham, NC 27710, USA, [1]Department of Human Genetics, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh, EH4 2XU, UK and [2]Department of Pediatrics, University of Vermont College of Medicine, Burlington, VT 05401, USA

Received May 8, 1995; Revised and Accepted July 10, 1995

Hereditary Hemorrhagic Telangiectasia or Osler–Rendu–Weber (ORW) disease is an autosomal dominant disorder characterized by multisystemic vascular dysplasia. Genetic linkage for some families was established to chromosome 9q34 (ORW1) (1,2) but genetic heterogeneity was demonstrated with the identification of some families clearly not linked to this locus (2–5). A predisposition for the development of pulmonary arteriovenous malformations (PAVMs) appears to segregate in ORW1 families (3–5). Recently, endoglin (ENG), a transforming growth factor β (TGF-β) binding protein, was identified as the ORW1 locus, with three independent mutations identified in affected individuals (6). We now report an additional six endoglin mutations including two in previously described ORW1 families. The type of mutations observed and their location in certain exons, with the apparent absence or under-representation of mutations in other exons, suggests a dominant-negative mechanism of receptor malfunctioning as the molecular basis of this disorder.

A panel of 80 DNA samples used in the initial mutation screen included affected individuals from nine families previously linked to 9q34 (1,3–5). Most of the remaining 71 samples were from families with a history of PAVM involvement, increasing the likelihood that each individual would harbor an independent endoglin mutation.

Primers were designed to amplify 11 of the 14 endoglin exons as previously described (6), although some primer pairs were redesigned to create smaller amplification products. The revised primer sequences and reaction conditions are available through GenBank Databank as entry G00-553-288. Heteroduplex analysis was performed and PCR products were directly sequenced as previously described (6). Table 1 catalogs the endoglin mutation summary to date, including three previously published mutations (6). Each of the mutations was seen only in a single chromosome in over 160 total chromosomes analyzed. Where additional family members were available, the mutations were shown to co-segregate with the disease phenotype.

Newly described mutations include a C to T base pair substitution creating a stop codon in exon 10 in a previously described ORW1 family (4). The normal and mutant bases were seen in direct sequencing of the PCR product (Fig. 1a). This point mutation removes a *Bbv*I restriction site. Exon 10 was amplified from all available family members and digested with *Bbv*I. All unaffected family members revealed the resulting digested fragments (normal allele), whereas all affec-

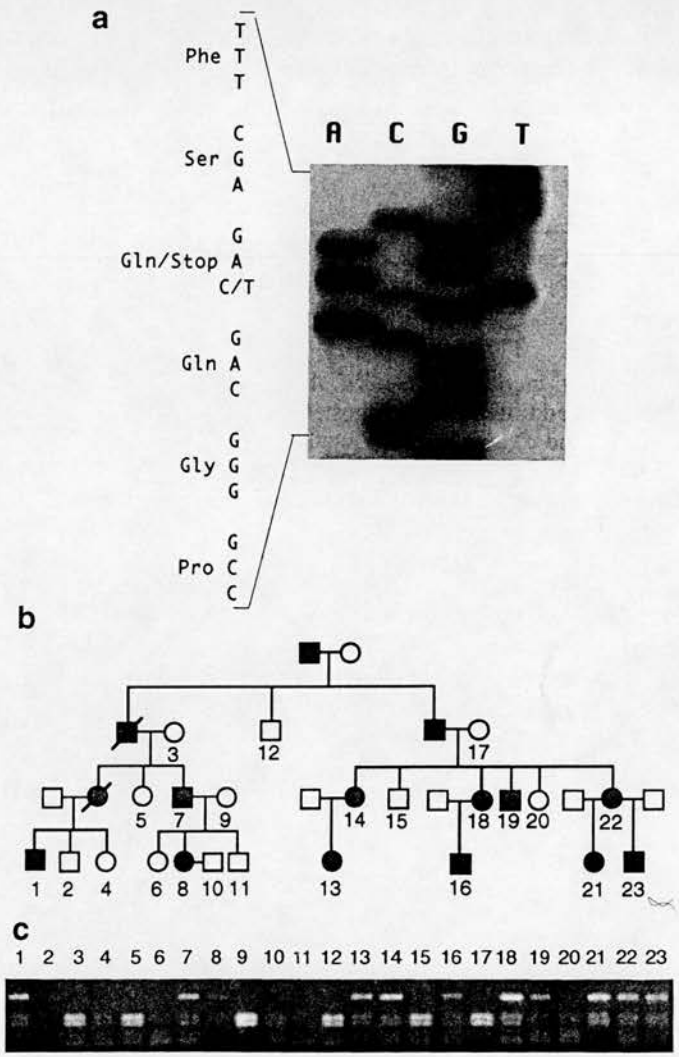


Figure 1. (**a**) Sequence of an exon 10 mutation in sample 1247 of family 32. Direct sequencing reveals both the normal (C) and the mutant (T) at nucleotide 1414, which creates a premature termination at codon 472, CAG to TAG. (**b**) Pedigree of family 32 denoting affected (closed symbols) and unaffected (open symbols) individuals. Numbers beneath the symbols refer to lane position in: (**c**) *Bbv*I digests of amplified exon 10 from family 32. This mutation removes a *Bbv*I site in the amplification product. All unaffected members reveal the digested doublet bands, whereas all affected individuals show the presence of three bands, corresponding to the uncut (mutant) and cut (normal) fragments of exon 10.

*To whom correspondence should be addressed

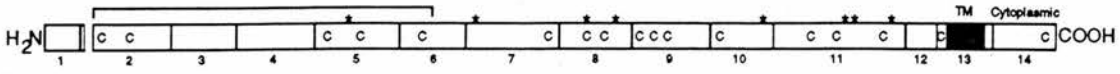


Figure 2. Truncation mutations in relation to the structure of endoglin. A schematic of the endoglin protein is shown to scale with contributions from each exon shown by the boxes numbered below. Amino acid residues from the first and a portion of the second exon are cleaved from the mature protein and are shown separated from the main body of the protein. Exons 2–12 encode the extracellular domain, exon 13 encodes the transmembrane (TM) domain and exon 14 encodes the cytoplasmic domain. The position of the cysteine residues are marked with C. The approximate positions of truncation mutations, whether nonsense or frameshifts, are shown with the asterisks. The bracket delineates the betaglycan-related amino-terminal domain of endoglin (8), which may correspond to the TGF-β binding domain (see text).

Table 1. Endoglin mutation summary

Exon	Sample no.	Family no.	Type of mutation
5	1152		Substitution at position 587*, creates stop codon TGG to TAG
7	8019	3186 (5)	Deletion (39 bp) at position 882 (6)
7	1159		Substitution at position 831, creates stop codon TAC to TA**G** (6)
8	1275	56 (4)	Substitution at position 1050, creates stop codon TGT to TG**A**
8	1160		Insertion (1 bp) at position 1111, frameshift creates premature stop codon
10	1247	32 (4)	Substitution at position 1414, creates stop codon **C**AG to **T**AG
11	2061		Deletion (2 bp) at position 1553, frameshift creates premature stop codon (6)
11	1162		Deletion (1 bp) at position 1655, frameshift creates premature stop codon
11	2241		Deletion (2 bp) at position 1550, frameshift creates premature stop codon

*Positions in endoglin as in reference 6.

ted family members revealed the uncut fragment (mutant allele) in addition to the digested fragments (normal allele) (Fig. 1b and c). Additional mutations are listed in Table 1.

Mutations have currently been identified in exons 5, 7, 8, 10 and 11, and all but one would produce a truncated protein if translated. We have not detected mutations in exons 2, 4, 6, 12, 13 and 14 using heteroduplex analysis and we have yet to develop PCR assays for exons 1, 3 and 9. The mutations observed thus far appear to cluster in a defined region when mapped to the endoglin protein (Fig. 2). For reference, exon 1 of the endoglin gene contains the cleaved signal peptide, exons 2–12 contain the extracellular domain of the mature endoglin, exon 13 contains the transmembrane domain and exon 14 contains the intracellular domain (6).

We had originally proposed a cellular-recessive model suggesting that endoglin mutations were loss-of-function alleles requiring somatic inactivation of the normal allele for vascular lesion development (6). With the apparent clustering of premature termination mutations observed in exons 5–11, we now propose a model involving the translation of a truncated protein product rather than a complete loss-of-function of the endoglin gene product. This dominant-negative model would suggest that the normal endoglin function is disrupted as a consequence of an effect caused by the mutant protein.

Endoglin is a homodimeric integral membrane glycoprotein expressed constitutively on vascular endothelial cells of capillaries, arterioles and venules (7). Endoglin binds TGF-β1 and TGF-β3 with high affinity (8). Transforming growth factor-β modulates several processes of endothelial cells including migration, proliferation, adhesion and extracellular matrix composition and organization (9). Perturbation of any one or a combination of these effects may lead to the development of the vascular lesions seen in ORW.

Endoglin is isolated from endothelial cells as a homodimer that can be dissociated by reducing agents (7,10). In affected individuals we postulate the assembly of normal/mutant heterodimers which do not function properly or which lead to rapid turnover of the receptor complex. Three cysteines are present in the processed endoglin protein prior to the first known truncation mutation in exon 5 (Fig. 2). One or more of these may be required for inter-molecular disulfide linkage, which might explain the apparent lack of truncation mutations prior to exon 5. Binding of TGF-β may also be required for the dominant-negative effect. The TGF-β binding domain of endoglin is not known, but this domain has been mapped in betaglycan (the type III TGF-β receptor), a related TGF-β binding protein. Deletion of the amino-terminal endoglin-related domain of betaglycan abolishes TGF-β binding activity (11).

Alternatively, the mutant endoglin allele might produce a soluble polypeptide which can bind and sequester the TGF-β ligand. Deletion mutants of betaglycan which remove the transmembrane and cytoplasmic domains produce soluble polypeptides which can bind TGF-β and inhibit TGF-β binding to membrane receptors (11). In the same way, soluble endoglin may act as an antagonist of TGF-β signaling.

For either of the proposed mechanisms, and in light of the observation that the vascular lesions of ORW are localized and discreet, we hypothesize that an initiating event is required to generate each lesion. This event could be endothelial wall damage due to thrombus formation or local inflammation. Endothelial cells having compromised endoglin function might respond poorly to TGF-β and form abnormal vessels or be deficient in repair of the vessel wall following such damage.

ACKNOWLEDGEMENTS

We gratefully acknowledge the individual patients and family members who participated in this study as well as the enthusiastic support of the Hereditary Hemorrhagic Telangiectasia Foundation International. This study was supported by NIH grant HL 49171 and a Baxter Scholar Award to D.A.M.

REFERENCES

1. McDonald,M.T., Papenberg,K.A., Ghosh,S. Glatfelter,A.A., Biesecker,B.B., Helmbold,E.A., Markel,D.S., Zolotor,A., McKinnon,W.C., Vanderstoep,J.L., Jackson,C.E., Iannuzzi,M., Collins,F.S., Boehnke,M., Porteous,M.E., Guttmacher,A.E. and Marchuk,D.A. (1994). A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33–34. *Nature Genet.*, **6**, 197–204 .
2. Shovlin,C., Hughes,J.M.B., Tuddenham,E.G.D., Temperley,I., Perembelon,Y.F.N., Scott,J., Seidman,C.E. and Seidman,J.G. (1994). A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genet.*, **6**, 205–209 .
3. Porteous,M.E.M., Curtis,A., Williams,O., Marchuk,D., Bhattacharya,S.S. and Burn,J. (1994). Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *J. Med. Genet.*, **31**, 925–926.
4. McAllister,K.A., Lennon,F., Bowles-Biesecker,B.,McKinnon,W.C., Helmbold,E.A., Markel,D.S., Jackson,C.E., Guttmacher,A.E., Pericak-Vance,M.A. and Marchuk,D.A. (1994). Genetic heterogeneity in hereditary hemorrhagic telangiectasia: possible correlation with clinical phenotype. *J. Med. Genet.*, **31**, 927–932.
5. Heutink,P., Haitjema,T., Breedveld,G.J., Janssen,B., Sandkuijl,L.A., Bontekoe,C.J.M., Westerman,C.J.J. and Oostra,B.A. (1994). Linkage of hereditary hemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. *J. Med. Genet.*, **31**, 933–936.
6. McAllister,K.A., Grogg,K.M., Johnson,D.W., Gallione,C.J., Baldwin,M.A., Jackson,C.E., Helmbold,E.A., Markel,D.S., McKinnon,W.C., Murrell,J., McCormick,M.K., Pericak-Vance,M.A., Heutink,P., Oostra,B.A., Haitjema,T., Westerman,C.J.J., Porteous,M.E., Guttmacher,A.E., Letarte,M. and Marchuk,D.A. (1994). Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genet.*, **8**, 345–351.
7. Gougos,A. and Letarte,M. (1990). Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J. Biol. Chem.*, **265**, 8361–8364.
8. Cheifetz,S., Bellon,T., Cales,C., Vera,S., Bernabeu,C., Massague,J. and Letarte,M. (1992). Endoglin is a component of the transforming growth factor-β receptor system in human endothelial cells. *J. Biol. Chem.*, **267**, 19 027–19 030.
9. Madri,J.A., Kocher,O., Merwin,J.R., Bell,L., Tucker,A. and Basson,C.T. (1989). Interactions of vascular cells with transforming growth factors-β. *Ann. NY Acad. Sci.*, **593**, 243–258.
10. Gougos,A. and Letarte,M. (1988). Biochemical characterization of the 44G4 antigen from the HOON Pre-B leukemic cell line. *J. Immunol.*, **141**, 1934–1940.
11. Lopez-Casillas,F., Payne,H.M., Andres,J.L. and Massague,J. (1994). Betaglycan can act as a dual modulator of TGF-β access to signaling receptors: mapping of ligand binding and GAG attachment sites. *J. Cell Biol.*, **124**, 557–568.

A Second Locus for Hereditary Hemorrhagic Telangiectasia Maps to Chromosome 12

David W. Johnson,[1] Jonathan N. Berg,[2] Carol J. Gallione,[1] Kimberly A. McAllister,[1] Jon P. Warner,[2] Elizabeth A. Helmbold,[3] Dorene S. Markel,[3] Charles E. Jackson,[3,4] Mary E.M. Porteous,[2] and Douglas A. Marchuk[1]

[1]Department of Genetics, Duke University Medical Center, Durham, North Carolina 27710; [2]Department of Human Genetics, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, United Kingdom; [3]Family Studies Core, Human Genome Center, University of Michigan Medical Center, Ann Arbor, Michigan 48109; [4]Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit, Michigan 48202

5:21–28 ISSN 1054-9803/95 $5.00

A Second Locus for Hereditary Hemorrhagic Telangiectasia Maps to Chromosome 12

David W. Johnson,[1] Jonathan N. Berg,[2] Carol J. Gallione,[1] Kimberly A. McAllister,[1] Jon P. Warner,[2] Elizabeth A. Helmbold,[3] Dorene S. Markel,[3] Charles E. Jackson,[3,4] Mary E.M. Porteous,[2] and Douglas A. Marchuk[1,5]

[1]Department of Genetics, Duke University Medical Center, Durham, North Carolina 27710; [2]Department of Human Genetics, University of Edinburgh, Western General Hospital, Edinburgh, EH4 2XU, United Kingdom; [3]Family Studies Core, Human Genome Center, University of Michigan Medical Center, Ann Arbor, Michigan 48109; [4]Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit, Michigan 48202

Hereditary hemorrhagic telangiectasia (HHT) or Osler-Rendu-Weber (ORW) disease is an autosomal dominant vascular dysplasia. Initial linkage studies identified an ORW gene localized to 9q33–q34 but with some families clearly excluding this region. A probable correlation in clinical phenotype between the 9q3-linked families and unlinked families was described with a significantly lower incidence of pulmonary arteriovenous malformations observed in the unlinked families. In this study we examined four unrelated ORW families for which linkage to chromosome 9q33–q34 has been previously excluded. Linkage was established for all four families to markers on chromosome 12, with a combined maximum lod score of 10.77 ($\theta = 0.04$) with D12S339. Mapping of crossovers using haplotype analysis indicated that the candidate region lies in an 11-cM interval between D12S345 and D12S339, in the pericentromeric region of chromosome 12. A map location for a second ORW locus is thus established that exhibits a significantly reduced incidence of pulmonary involvement.

Hereditary hemorrhagic telangiectasia (HHT) or Osler-Rendu-Weber (ORW) disease is an autosomal dominant multisystemic vascular dysplasia. The three primary types of angiodysplasia exhibited are telangiectases (mucosal, dermal, and visceral), arterial venous malformations (AVMs, particularly pulmonary, cerebral, and hepatic), and aneurysms. Penetrance is age dependent but is considered nearly complete by age 40 (Plauchu et al. 1989; Porteous et al. 1992).

The most common clinical feature of ORW is recurrent epistaxis from vascular lesions in the nasal epithelium, which affects ~90% of ORW patients (Aassar et al. 1991). Cutaneous telangiectases are seen in ~70% of affected individuals (Brant et al. 1989; Plauchu et al. 1989). Gastrointestinal bleeding occurs in ~20% of ORW patients (Reilly and Nostrant 1984; Vase and Grove 1986). Pulmonary arteriovenous malformations (PAVMs) occur in ~20% of ORW patients (Dines et al. 1974) and are associated with serious complications including stroke and brain abscess (White and Pollak 1994). Neurologic manifestations include cerebral arteriovenous malformation, aneurysm, and migraine headache (Willinsky et al. 1990).

Incidence rates have been calculated for the spectrum of symptoms in a number of retrospective and prospective studies (Reilly and Nostrant 1984; Peery 1987; Plauchu et al. 1989; Porteous et al. 1992). Great variability of expression and severity is seen even among members of the same family, indicating that factors in addition to the inherited germ-line mutation contribute to the individual's phenotype. Until recently it was unclear whether all families present the entire spectrum of clinical features. Differences might be attributable to either allelic or locus heterogeneity.

Genetic linkage for some ORW families was established with markers on 9q3 (McDonald et al. 1994; Shovlin et al. 1994). This locus (ORW1) was

[5]Corresponding author.
E-MAIL march004@mc.duke.edu; FAX (919) 684-2790.

subsequently shown to be the endoglin gene that encodes a transforming growth factor-β (TGF-β)-binding protein of endothelial cells (McAllister et al. 1994b). Locus heterogeneity was indicated by families that excluded linkage to this region (Heutink et al. 1994; McAllister et al. 1994a; Porteous et al. 1994; Shovlin et al. 1994). Several reports (Heutink et al. 1994; McAllister et al. 1994a; Porteous et al. 1994) have documented a significantly lower incidence (in some cases, an absence) of PAVMs in non-9q-linked families when compared to 9q-linked families (ORW1). In this report we show that at least some of the genetic and clinical heterogeneity can be explained by a second ORW locus mapping to chromosome 12.

RESULTS

Our approach to establish genetic linkage with these families included a candidate gene approach, based on the known map positions of ~20 growth factors (and their receptors), mitogenic or growth inhibitory for endothelial cells. In addition to these loci, members of the TGF-β receptor superfamily, including the activin receptors and activin receptor-like kinases, were considered candidate genes based on the identification of endoglin, a TGF-β-binding protein, as the gene for ORW1 (McAllister et al. 1994b).

Linkage analysis revealed that the disease gene in four families (Fig. 1) is linked to markers mapping to the pericentromeric region of chromosome 12. The pairwise lod scores between the disease gene and these markers are given in Table 1. Markers D12S85 and D12S339 showed the highest maximum combined lod scores of 9.06 ($\theta = 0$) and 10.77 ($\theta = 0.04$), respectively. Marker informativity varied considerably among the four families, causing large differences in lod scores for closely spaced markers. For example, in family 67, marker D12S368 showed a much lower lod score (maximum 0.21 at $\theta = 0.21$) than marker D12S339 (maximum 3.94 at $\theta = 0.05$) which is only 3 cM distal, owing primarily to uninformative genotypes for key affected parents 67-III-10 and 67-III-19. Each family showed highly positive lod scores with at least one of the markers in this region (Table 1).

Informative recombinant individuals were identified to define more precisely the candidate interval. Haplotypes were determined across a 37-cM interval for each family, but crossovers were scored only from affected individuals, because of potential nonpenetrance of unaffected members. There were no unambiguous crossovers in the disease haplotypes in families 2 and 17. Affected individuals from two families (33-IV-1 and 67-IV-5) showed crossovers between markers D12S310 and D12S333 (Fig. 2). An affected individual in family 67 (67-IV-2) showed a crossover between D12S345 and D12S85. These place the disease locus in the region between D12S345 and 12qter. One affected individual in family 33 (33-III-3; Fig. 1) showed a crossover between markers D12S339 and D12S85 that provided the distal boundary for the disease locus. The candidate interval is thus bordered by D12S345 and D12S339, an interval of 11 cM (Fig. 2).

Several individuals diagnosed as unaffected (17-IV-3, age 20; 33-III-5, age 31; 33-III-9, age 27; 67-III-2, age 45; 67-IV-7, age 40; 67-IV-10, age 34; and 67-IV-11, age 49) showed the disease haplotype through most or all of the region, including the entire 11-cM candidate interval. These may represent complex recombinants or, more likely, nonpenetrant individuals. The four individuals from family 67 caused the maximum lod score to occur at a $\theta = 0.13$, rather than at zero with marker D12S85.

Significantly, no individual diagnosed as affected from any of these families shows an unaffected (normal) haplotype. Diagnoses in our non-9q3-linked families were often difficult, perhaps because of lower penetrance or expressivity of the ORW2 locus compared with ORW1. However, only those individuals meeting the previously defined stringent diagnostic criteria were assigned an affected phenotype. From our observations we believe there also may be a delayed age of onset of symptoms in ORW2, but the limited data do not justify gross revision of the age-dependent penetrance model used previously. Nonetheless, when the model was changed from maximum penetrance of 97% to 75% (age 40 and above), evidence for linkage of ORW2 to the three most closely linked markers remained strong (Table 2).

DISCUSSION

These data provide strong evidence for the existence of an ORW locus mapping to chromosome 12. The centromere has been localized on the physical map between markers D12S59 and D12S85 (Kucherlapati et al. 1994). Based on a comparison of physical and genetic maps, D12S59 is located within the interval between D12S345 and D12S85, an area contained within

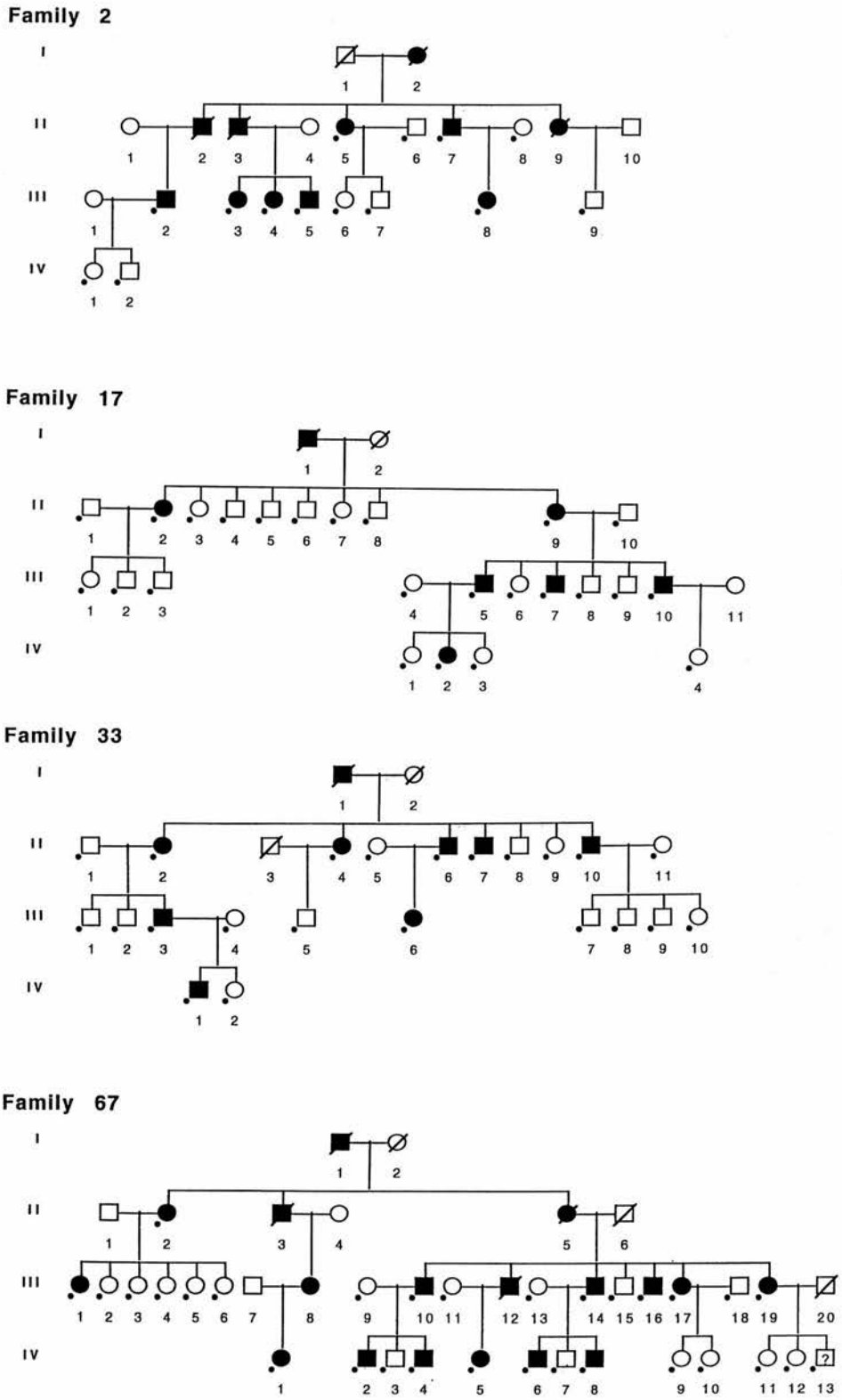


Figure 1 Pedigree drawings of the four ORW families, using standard symbols. A dot below a symbol indicates an individual who was sampled for this study.

the candidate region for ORW2. Therefore, we provisionally assign the ORW2 locus as pericentromeric, as our data do not allow definitive assignment to a particular arm of chromosome 12.

Previously, we had suggested the possibility of an ORW locus mapping to chromosome 3p22 (McAllister et al. 1994b), based on a two-point lod score of 1.43 in family 33 with marker D3S1211. This marker maps near the TGF-β receptor II, which was considered a candidate for this locus. The stronger statistical support for linkage of the disease locus in family 33 to chromosome 12 (lod score of 2.64 at $\theta = 0$ for marker D12S85, Table 1) suggests that the actual location for the ORW locus in this family is the pericentromeric region of chromosome 12. No additional families have shown linkage to this region of chromosome 3. Thus, we have no evidence for an ORW locus mapping to 3p22. Likewise we have no evidence for other ORW loci, as we have no families that exclude both the ORW1 locus on 9q and the ORW2 locus on chromosome 12.

The absence of PAVMs in ORW family 67 (first described in this report) supports the previously published observation of a

Table 1. Pairwise lod scores between HHT and chromosome 12 marker loci

		LOD scores (Z) at different recombination fractions θ							
Locus	Family	0.01	0.05	0.10	0.20	0.30	0.40	Zmax	θ
D12S310	2	−0.90	−0.55	−0.33	−0.13	−0.06	−0.02	−0.02	0.40
	17	2.12	1.97	1.77	1.35	0.88	0.37	2.15	0.00
	33	−1.59	−0.37	0.01	0.18	0.12	0.03	0.18	0.20
	67	−1.18	−0.38	0.01	0.29	0.28	0.15	0.31	0.24
	Total	−1.55	0.66	1.46	1.70	1.23	0.53	1.74	0.17
D12S333	2	−1.10	−0.46	−0.24	−0.09	−0.04	−0.02	−0.02	0.40
	17	3.88	3.65	3.33	2.59	1.71	0.69	3.94	0.00
	33	0.63	0.81	0.86	0.76	0.52	0.22	0.86	0.10
	67	−2.92	−0.76	0.05	0.49	0.37	0.05	0.50	0.22
	Total	0.50	3.24	4.00	3.75	2.55	0.94	4.08	0.13
D12S345	2	−0.97	0.22	0.56	0.63	0.43	0.14	0.65	0.16
	17	3.55	3.30	2.96	2.24	1.43	0.56	3.61	0.00
	33	1.20	1.36	1.37	1.15	0.74	0.27	1.38	0.08
	67	−2.09	0.27	1.16	1.59	1.31	0.66	1.59	0.20
	Total	1.69	5.15	6.06	5.61	3.91	1.63	6.13	0.12
D12S85	2	1.48	1.33	1.14	0.78	0.43	0.13	1.51	0.00
	17	4.24	3.96	3.59	2.77	1.82	0.74	4.31	0.00
	33	2.60	2.44	2.22	1.70	1.08	0.39	2.64	0.00
	67	0.71	0.93	1.02	0.97	0.75	0.42	1.02	0.13
	Total	9.02	8.66	7.98	6.22	4.08	1.69	9.06	0.00
D12S339	2	2.16	1.97	1.73	1.22	0.70	0.23	2.21	0.00
	17	3.88	3.65	3.33	2.59	1.71	0.69	3.94	0.00
	33	0.60	1.16	1.27	1.10	0.72	0.26	1.27	0.10
	67	3.75	3.94	3.84	3.18	2.21	1.02	3.94	0.05
	Total	10.39	10.73	10.16	8.09	5.34	2.20	10.77	0.04
D12S368	2	1.98	1.81	1.59	1.12	0.64	0.21	2.02	0.00
	17	2.35	2.42	2.32	1.88	1.24	0.47	2.42	0.04
	33	1.15	1.66	1.70	1.41	0.91	0.33	1.71	0.08
	67	−0.70	−0.20	0.07	0.21	0.17	0.07	0.21	0.21
	Total	4.78	5.69	5.68	4.63	2.96	1.07	5.75	0.07
D12S359	2	0.35	0.86	0.92	0.74	0.43	0.11	0.92	0.09
	17	2.35	2.42	2.32	1.88	1.24	0.46	2.42	0.04
	33	−1.18	−0.53	−0.28	−0.07	0.02	0.05	0.05	0.40
	67	1.37	1.94	2.14	1.94	1.35	0.51	2.15	0.11
	Total	2.89	4.69	5.11	4.49	3.03	1.13	5.11	0.11
D12S355	2	0.30	0.81	0.88	0.70	0.40	0.12	0.88	0.09
	17	2.16	2.24	2.16	1.75	1.15	0.42	2.24	0.05
	33	0.81	1.34	1.41	1.18	0.76	0.28	1.41	0.09
	67	3.21	3.49	3.48	2.99	2.14	1.04	3.52	0.07
	Total	6.48	7.88	7.93	6.63	4.46	1.85	8.00	0.08

Age-dependent penetrance of HHT was estimated using the following three liability classes: ages 0–20 (penetrance 0.2425); ages 21–40 (penetrance 0.7275); ages >40 (penetrance 0.9700).

much reduced incidence of PAVMs in other non-ORW1 families (Heutink et al. 1994; McAllister et al. 1994a; Porteous et al. 1994). Additional ORW1 and ORW2 families must be studied before accurate risk levels for the development of PAVMs can be calculated.

With the establishment of a second locus for ORW on chromosome 12, further studies are now justified comparing both the incidence and severity of the clinical features between ORW1 and ORW2 families. These might provide vital information resulting in more accurate counseling for disease prognosis and more effective management of affected individuals.

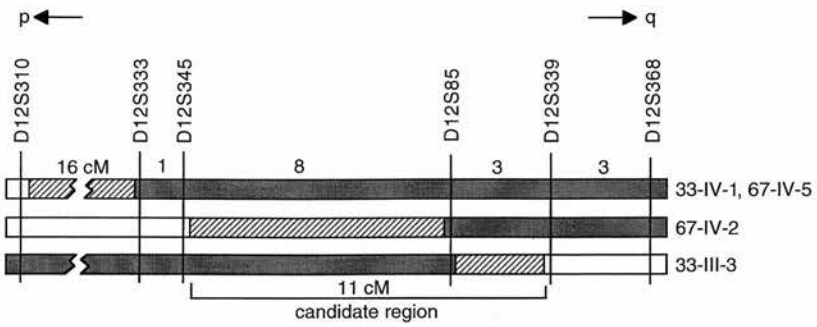


Figure 2 Graphic representation of the recombinant chromosomes identified in ORW families that define the candidate interval. The hatched boxes represent the minimum region where each recombination occurred, and the dark boxes represent the chromosome 12 region that segregates with the disease. The individuals harboring the recombinants are indicated by a code consisting of the family number, the generation number and, finally, the person number, each of which corresponds to Fig. 1. Map distances are based on Gyapay et al. (1994). The centromere is located between D12S345 and D12S85, based on a comparison of genetic and physical maps.

The identification of endoglin as the gene mutated in ORW1 (McAllister et al. 1994b) indicates the importance of TGF-β signaling pathways in the pathology of this form of the disease. The ORW2 gene may encode an endothelial cell receptor involved in a similar process of cell signaling. Two potential candidate genes, ACVRLK1 and ACVRLK4 (activin receptor-like kinases), map to chromosome 12q11–q14 based on hybridization to a chromosome 12 hybrid mapping panel (Kucherlapati et al. 1994). ACVRLK1 (ALK1) encodes a putative cell-surface receptor with a serine/threonine kinase domain, expressed predominantly from endothelial cells (Attisano et al. 1993). It can associate with the type II TGF-β receptor after overexpression in COS cells, although its ligand in vivo remains unknown (ten Dijke et al. 1994). ACVRLK4 (ALK4 or SKR2) encodes a ubiquitously expressed cell-surface receptor with a serine/threonine kinase domain (ten Dijke et al. 1993) that binds the growth factor activin (ten Dijke et al. 1994). Growth of vascular endothelial cells in culture is inhibited by activin-A and TGF-β causes an additive inhibitory effect (McCarthy and Bicknell 1993).

A third candidate gene that is under consideration is ITGA5, encoding the integrin α5 subunit, which together with the integrin β1 subunit forms the fibronectin receptor. The ITGA5 locus

Table 2. Pairwise lod scores between HHT and chromosome 12 marker loci, using a revised penetrance model

		Lod scores (Z) at different recombination fractions (θ)							
Locus	**Family**	**0.01**	**0.05**	**0.10**	**0.20**	**0.30**	**0.40**	**Zmax**	**θ**
D12S345	2	−1.22	−0.01	0.35	0.46	0.31	0.09	0.47	0.17
	17	2.97	2.74	2.45	1.81	1.12	0.43	3.03	0.00
	33	1.76	1.68	1.55	1.19	0.74	0.26	1.78	0.00
	67	−1.21	0.60	1.22	1.44	1.12	0.54	1.45	0.18
	Total	2.30	5.01	5.57	4.90	3.29	1.31	5.57	0.11
D12S85	2	1.24	1.10	0.93	0.61	0.32	0.09	1.27	0.00
	17	3.41	3.17	2.86	2.16	1.37	0.53	3.47	0.00
	33	2.35	2.21	2.00	1.52	0.95	0.34	2.38	0.00
	67	1.22	1.20	1.14	0.97	0.71	0.38	1.23	0.00
	Total	8.22	7.68	6.93	5.26	3.34	1.34	8.35	0.00
D12S339	2	1.91	1.74	1.51	1.04	1.57	0.17	1.96	0.00
	17	3.05	2.86	2.59	1.98	1.26	0.47	3.10	0.00
	33	0.35	0.93	1.05	0.92	0.59	0.22	1.05	0.11
	67	4.54	4.27	3.89	3.03	2.02	0.09	4.60	0.00
	Total	9.86	9.79	9.04	6.97	4.45	1.76	9.98	0.02

Age-dependent penetrance of HHT was estimated using the following three liability classes: ages 0–20 (penetrance 0.2425); ages 21–40 (penetrance 0.7275); ages >40 (penetrance 0.7500).

has been mapped cytogenetically to 12q11–q13 (Sosnoski et al. 1988). Integrin α5 is expressed in cell types including endothelial cells, and its expression in endothelial cells is modulated by TGF-β (Enenstein et al. 1992). The α5β1 fibronectin receptor can be detected at the site of intercellular contact between endothelial cells where it regulates the integrity and permeability of the monolayer (Lampugnani et al. 1991). In addition, extracellular matrix components, including fibronectin mediate complex changes in the migratory behavior of endothelial cells (Hauser et al. 1993).

METHODS

Clinical Evaluation

Family members were interviewed and examined by a medical geneticist. The interview included relevant medical history information such as age of onset and pattern of any epistaxis, evidence for the presence of pulmonary or cerebral arteriovenous malformations, and history of gastrointestinal or pulmonary bleeding. Pertinent medical records, including chest angiography, were obtained and reviewed when available. Individuals presenting with dyspnea or cyanosis were referred for radiological examination for PAVMs. Family history was reviewed. Physical examination included a search for telangiectases on the lips, tongue, oral or nasal mucosa, cheeks, fingers, or elsewhere. Affected status was assigned based on the presence of at least two of the following three criteria: telangiectasia, recurrent epistaxis, and an affected first-degree relative (McAllister et al. 1994a; McDonald et al. 1994). Informed consent was obtained and blood samples were drawn for DNA extraction, using standard methods.

Family Description

Four families (Fig. 1) for which linkage to the ORW1 locus was excluded, were included in this study; family 2 (Porteous et al. 1994), families 17 and 33 (McAllister et al. 1994a), and family 67 (D.W. Johnson and D.A. Marchuk, unpubl.). There were no affected individuals with clinical evidence for PAVMs, despite attempts at uncovering potentially undiagnosed lesions. It appeared that the age of onset of symptoms in these families was later than in the 9q-linked families described previously (McAllister et al. 1994a; McDonald et al. 1994). No other clear clinical distinctions could be made between these and the previously described ORW1 families.

Family 2 consists of 14 sampled individuals including 7 affected members (Porteous et al. 1994). There was a wide range of severity of epistaxes and telangiectases among affected individuals. Pulse oximetry and chest X-ray showed no evidence of PAVMs. The pedigree drawings for this and the following families are shown in Figure 1.

Family 17 consists of 24 sampled individuals, including 6 affected members (McAllister et al. 1994a). Frequent epistaxis and telangiectases are also prevalent in this family. No family member has ever been diagnosed with a PAVM.

Family 33 consists of 22 sampled individuals with 8 affected members (McAllister et al. 1994a). Symptoms in the majority of affected members include epistaxis and mucocutaneous telangiectases. One affected individual was diagnosed as having a cerebral AVM, and one member died of hepatic failure. No affected individuals had any evidence of PAVMs.

Family 67 is a new kindred consisting of 30 sampled individuals with 13 affected members. Epistaxis and mucocutaneous telangiectases are the common clinical features seen. One affected individual had a cerebral hemorrhage, and one had a history of gastrointestinal hemorrhage. Again, no affected individual had any evidence of PAVMs.

Genotyping and Linkage Analysis

Analyses of simple sequence repeat markers was performed as described previously (McDonald et al. 1994). Two-point linkage analysis was also performed as described previously (McAllister et al. 1994a). Briefly, ORW was analyzed as an autosomal dominant disorder with age-dependent penetrance. For estimating the age-dependent penetrance, three liability classes were created with age ranges 0–20 (penetrance 0.2425), 21–40 (0.7275), and >40 (0.9700). For the data presented in Table 2, penetrance for the final age group was set at 0.7500. Disease frequency was set at 0.0001. Two-point linkage analysis was performed on a Sun Sparc station 10 using the MLINK subprogram of the LINKAGE computer package (v. 4.9; Lathrop et al. 1984). Allele frequency estimates for the markers were obtained from Genome Data Base.

NOTE ADDED IN PROOF

While this manuscript was under review a separate report of linkage of an HHT locus to chromosome 12 was published (Vincent et al. 1995. Hum. Molec. Genet. **4:** 945–949).

ACKNOWLEDGMENTS

We gratefully acknowledge the family members who participated in this study, as well as the enthusiastic support of the Hereditary Hemorrhagic Telangiectasia Foundation, International. We thank Dr. Robert I. White (Yale University Medical School) for alerting us to family 67, and Michele Walsh of the General Clinical Research Center, University of Michigan Medical School, for cell immortalization. This study was supported by National Institutes of Health (grant HL 49171) and a Baxter Foundation award to D.A.M.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Aassar, O.S., C.M. Friedman, and R.I. White Jr. 1991. The natural history of epistaxis in hereditary hemorrhagic telangiectasia. *Laryngoscope* **101:** 977–980.

Attisano, L., J. Carcamo, F. Ventura, F.M.B. Weis, J. Massague, and J.L. Wrana. 1993. Identification of human activin and TGF-β type I receptors that form heteromeric kinase complexes with type II receptors. *Cell* **75:** 671–680.

Brant, A.M., A.P. Schachat, and R.I. White. 1989. Ocular manifestations in hereditary hemorrhagic telangiectasia (Rendu-Osler-Weber disease). *Am. J. Ophthalmol.* **107:** 642–646.

Dines, D.E., R.A. Arms, P.E. Bernatz, and M.R. Gomes. 1974. Pulmonary arteriovenous fistulas. *Mayo Clin. Proc.* **49:** 460–465.

Enenstein, J., N.S. Waleh, and R.H. Kramer. 1992. Basic FGF and TGF-beta differentially modulate integrin expression of human microvascular endothelial cells. *Exp. Cell Res.* **203:** 499–503.

Gyapay, G., J. Morissette, A. Vignal, C. Dib, C. Fizames, P. Millasseau, S. Marc, G. Bernardi, M. Lathrop, and J. Weissenbach. 1994. The 1993-94 Genethon human genetic linkage map. *Nature Genet.* **7:** 246–339.

Hauser, I.A., E. Setter, L. Bell, and J.A. Madri. 1993. Fibronectin expression correlates with U937 cell adhesion to migrating bovine aortic endothelial cells in vitro. *Am. J. Pathol.* **143:** 173–180.

Heutink, P., T. Haitjema, G.J. Breedveld, B. Janssen, L.A. Sandkuijl, C.J.M. Bontekoe, C.J.J. Westerman, and B. Oostra. 1994. Linkage of hereditary hemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. *J. Med. Genet.* **31:** 933–936.

Kucherlapati, R., I. Craig, and P. Marynen. 1994. Report of the second international workshop on human chromosome 12 mapping 1994. *Cytogenet. Cell Genet.* **67:** 246–264.

Lampugnani, M.G., M. Resnati, E. Dejana, and P.C. Marchisio. 1991. The role of integrins in the maintenance of endothelial monolayer integrity. *J. Cell Biol.* **112:** 479–490.

Lathrop, G.M., J.M. Lalouel, C. Julier, and J. Ott. 1984. Strategies for multilocus linkage analysis in humans. *Proc. Natl. Acad. Sci.* **81:** 3443–3446.

McAllister, K.A., F. Lennon, B. Bowles-Biesecker, W.C. McKinnon, E.A. Helmbold, D.S. Markel, C.E. Jackson, A.E. Guttmacher, M.A. Pericak-Vance, and D.A. Marchuk. 1994a. Genetic heterogeneity in hereditary hemorrhagic telangiectasia: Possible correlation with clinical phenotype. *J. Med. Genet.* **31:** 927–932.

McAllister, K.A., K.M. Grogg, D.W. Johnson, C.J. Gallione, M.A. Baldwin, C.E. Jackson, E.A. Helmbold, D.S. Markel, W.C. McKinnon, J. Murrell, M.K. McCormick, M.A. Pericak-Vance, B.A. Oostra, T. Haitjema, C.J.J. Westerman, M.E. Porteous, A.E. Guttmacher, M. Letarte, and D.A. Marchuk. 1994b. Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genet.* **8:** 345–351.

McCarthy, S.A., and R.Bicknell. 1993. Inhibition of vascular endothelial cell growth by activin-A. *J. Biol. Chem.* **268:** 23066–23071.

McDonald, M.T., K.A. Papenberg, S. Ghosh, A.A. Glatfelter, B.B. Biesecker, E.A. Helmbold, D.S. Markel, A. Zolotor, W.C. McKinnon, J.L. Vanderstoep, C.E. Jackson, M. Iannuzzi, F.S. Collins, M. Boehnke, M.E. Porteous, A.E. Guttmacher, and D.A. Marchuk. 1994. A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. *Nature Genet.* **6:** 197–204.

Peery, W.H. 1987. Clinical spectrum of hereditary hemorrhagic telangiectasia (Osler-Weber-Rendu disease). *Am. J. Medicine* **82:** 989–996.

Plauchu, H., J.-P. Chadarevian, A. Bideau, and J.-M. Robert. 1989. Age-related clinical profile of hereditary hemorrhagic telangiectasia in an epidemiologically recruited population. *Am. J. Med. Genet.* **32:** 291–297.

Porteous, M.E.M., J. Burn, and S.J. Proctor. 1992. Hereditary haemorrhagic telangiectasia: A clinical analysis. *J. Med. Genet.* **29:** 527–530.

Porteous, M.E.M., A. Curtis, O. Williams, D. Marchuk, S.S. Bhattacharya, and J. Burn. 1994. Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *J. Med. Genet.* **31:** 925–926.

Reilly, P.J., and T.T. Nostrant. 1984. Clinical manifestations of hereditary hemorrhagic telangiectasia. *Am. J. Gastroenterol.* **79:** 363–367.

Shovlin, C.L., J.M.B. Hughes, E.G.D. Tuddenham, I. Temperley, Y.F.N. Perembelon, J. Scott, C.E. Seidman, and J.G. Seidman. 1994. A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genet.* **6:** 205–209.

Sosnoski, D., B.S. Emanuel, A.L. Hawkins, P. van Tuinen, D.H. Ledbetter, R.L. Nussbaum, F.-T. Kaos, E. Schwartz, D. Phillips, J.S. Bennett, L.A. Fitzgerald, and M. Poncz. 1988. Chromosomal localization of the genes for the vitronectin and fibronectin receptors alpha-subunits and for platelet glycoproteins IIb and IIIa. *J. Clin. Invest.* **81:** 1993–1998.

ten Dijke, P., H. Ichijo, P. Franzen, P. Schulz, J. Saras, H. Toyoshima, C.-H. Heldin, and K. Miyazono. 1993. Activin receptor-like kinases: A novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. *Oncogene* **8:** 2879–2887.

ten Dijke, P., H. Yamashita, H. Ichijo, P. Franzen, M. Laiho, K. Miyazono, and C.-H. Heldin. 1994.

Characterization of type I receptors for transforming growth factor-β and activin. *Science* **264:** 101–104.

Vase, P., and O. Grove. 1986. Gastrointestinal lesions in hereditary hemorrhagic telangiectasia. *Gastroenterology* **91:** 1079–1083.

White, R.I., and J.S. Pollak 1994. Pulmonary arteriovenous malformations: options for management. *Ann. Thorac. Surg.* **57:** 519–521.

Willinsky, R.A., P. Lasjaunias, K. Terbrugge, and P. Burrows. 1990. Multiple cerebral arteriovenous malformations (AVMs). Review of our experience from 203 patients with cerebral vascular lesions. *Neuroradiology* **32:** 207–210.

Received May 3, 1995; accepted in revised form June 15, 1995.

Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2

D.W. Johnson[1], J.N. Berg[1,2], M.A. Baldwin[1], C.J. Gallione[1], I. Marondel[3], S.-J. Yoon[3], T.T. Stenzel[1,4], M. Speer[5], M.A. Pericak-Vance[5], A. Diamond[6], A.E. Guttmacher[7], C.E. Jackson[8], L. Attisano[9], R. Kucherlapati[3], M.E.M. Porteous[2] & D.A. Marchuk[1]

[1]*Department of Genetics, Duke University Medical Center, Box 3175, Durham, North Carolina 27710, USA*
[2]*Department of Human Genetics, University of Edinburgh, Edinburgh EH4 2XU, UK*
[3]*Department of Molecular Genetics, Albert Einstein College of Medicine, 1300 Morris Park Ave., Bronx, New York 10461, USA*
[4]*Department of Pathology, Duke University Medical Center, Durham, North Carolina 27710, USA*
[5]*Department of Medicine, Division of Neurology, Duke University Medical Center, Durham, North Carolina 27710, USA*
[6]*Department of Immunology, University of Newcastle upon Tyne, Newcastle upon Tyne NE2 4AA, UK*
[7]*Department of Pediatrics, University of Vermont, Burlington, VT 05401, USA*
[8]*Department of Medicine, Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit, Michigan 48202, USA*
[9]*Department of Anatomy and Cell Biology, The University of Toronto, Toronto, Ontario M5S 1A8, Canada*

Correspondence should be addressed to D.A.M.

Hereditary haemorrhagic telangiectasia, or Osler-Rendu-Weber (ORW) syndrome, is an autosomal dominant vascular dysplasia. So far, two loci have been demonstrated for ORW. Linkage studies established an *ORW* locus at chromosome 9q3; endoglin was subsequently identified as the *ORW1* gene. A second locus, designated *ORW2,* was mapped to chromosome 12. Here we report a new 4 cM interval for *ORW2* that does not overlap with any previously defined. A 1.38-Mb YAC contig spans the entire interval. It includes the activin receptor like kinase 1 gene (*ACVRLK1* or *ALK1*), a member of the serine-threonine kinase receptor family expressed in endothelium. We report three mutations in the coding sequence of the *ALK1* gene in those families which show linkage of the ORW phenotype to chromosome 12. Our data suggest a critical role for *ALK1* in the control of blood vessel development or repair.

Hereditary haemorrhagic telangiectasia, or Osler-Rendu-Weber (ORW) syndrome, is an autosomal dominant multisystemic vascular dysplasia. It has an estimated prevalence of 1 in 40,000, although this may be an underestimate[1]. ORW is characterized by recurrent epistaxis, muco-cutaneous telangiectases, gastro-intestinal haemorrhage, and pulmonary, cerebral and hepatic arteriovenous malformations; all secondary manifestations of the underlying vascular dysplasia[2]. In general, affected individuals from most ORW families present with an overlapping constellation of clinical symptoms. Both variable expression and pleiotropy of ORW can cause diagnostic uncertainty, particularly in younger individuals, as many of the disease features become more pronounced with age.

Linkage studies have shown at least two distinct loci for ORW. The first, *ORW1,* was located on chromosome 9q3 and subsequently shown to be the endoglin gene[3–5]. Endoglin is a TGF-β binding protein expressed predominantly by endothelial cells and placenta[6,7]. Most mutations in endoglin appear to be protein truncations, suggesting a dominant-negative effect of binding protein function[8].

Locus heterogeneity for ORW was indicated by families which excluded linkage to endoglin[4,9–11]. Several reports[9–12] have documented a significantly higher prevalence of pulmonary arteriovenous malformations in ORW1 families when compared to non-ORW1. A second ORW locus (*ORW2*) was subsequently identified in the pericentromeric region of chromosome 12 (refs 13, 14). ORW2 may exhibit reduced penetrance[14], as a number of unaffected adults in ORW2 families show the disease haplotype over a broad region of chromosome 12.

We report further family assessment data and new crossovers that narrow the candidate interval for *ORW2* to a 4-cM region on chromosome 12q. The new candidate interval, which does not overlap that defined previously, is based on the re-evaluation of the disease status of the individual exhibiting a critical distal crossover. A physical map over the region demonstrates that the activin receptor like kinase 1 (also known as *ALK1, ACVRLK1* or *TSR1*) gene maps within the *ORW2* interval.

ALK1 is a type I cell-surface receptor for the TGF-β superfamily of ligands. It shares with other type I receptors a high degree of similarity in serine-threonine kinase subdomains, a glycine- and serine-rich region (called the GS domain) preceding the kinase domain, and the short C-terminal tail[15,16]. The ALK1 protein can associate with the TGF-β or activin type II receptors after cotransfection in COS cells, with the complex binding TGF-β or activin respectively[15,17,18]. However the ALK1 ligand *in vivo* remains unknown[16], due to the promiscuity of ligand binding of type I receptors in cell culture overexpression systems[19]. The high level of expression of ALK1 in endothelial cells and other highly vascularized tissues (lung and placenta) and lower expression levels in other cells and tissues[15] parallels that of endoglin. We report mutations in the coding region of the *ALK1* gene in ORW2 patients.

Genetic and physical map of the *ORW2* interval

We originally reported an 11-cM candidate interval for *ORW2* in the pericentromeric region of chromosome 12, between *D12S345* and *D12S339* (Fig. 1)[14]. New

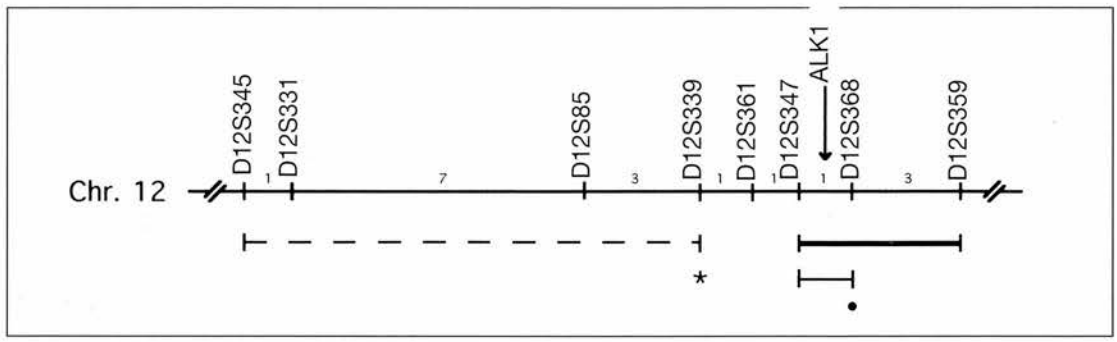


Fig. 1 Candidate intervals for *ORW2* based on crossovers in chromosome 12-linked ORW families. A genetic map of a portion of chromosome 12 is shown, with distances between markers listed in cM[20]. A published 11-cM interval is shown as a dashed line[14]. The distal crossover at *D12S339* was re-examined due to uncertainty of map location of the marker and re-assignment of disease status of the individual (see text). The remaining crossovers in clearly affected individuals give a 4-cM candidate interval between *D12S347* and *D12S359*, shown as a bold line. The inclusion of a crossover in an unaffected adult, shown with a dot, gives a 1-cM candidate interval between *D12S347* and *D12S368*.

ORW families were typed with markers in this region. Family 3 showed a maximum pairwise lod score of 4.14 ($\theta = 0$) with marker *D12S297*. Family 40 showed a maximum pairwise lod score of 2.42 ($\theta = 0$) with marker *D12S297*, close to its maximum possible score calculated under linkage simulation. The candidate interval as defined by family 3 does not overlap with the original region with regard to the distal boundary defined by an individual from a previous ORW2 family. This boundary was defined by marker *D12S339*, which shows a discordant position between the genetic[20] and physical maps (I.M. and R.K., unpublished observations).

In view of these apparent discrepancies, we elected to reassess the clinical data from the discordant critical individual from family 33 and to disregard data from *D12S339* in haplotype analysis. The distal crossover individual from family 33 (ref. 14) has no epistaxis or characteristic telangiectases; however, he was previously designated as affected on the basis of both a suspect lesion which had been excised from his foot, and his son's epistaxis. We subsequently obtained the micro-pathology report of the father's lesion, which was a pyogenic granuloma, not a characteristic ORW lesion. The unreliability of epistaxis as a diagnostic criterion was previously noted[4]. The diagnosis of both individuals was thus designated uncertain.

A crossover in an affected individual from family 3 (III-11) sets the new proximal boundary at marker *D12S347* (Fig. 2). An affected individual in family 67 (V-2) sets the new distal boundary at *D12S359* (Fig. 3), giving a critical 4 cM interval for *ORW2* (Fig. 1). A crossover in an unaffected individual (67-V-8; Fig. 3) would place the distal boundary for *ORW2* at *D12S368*, and would locate *ORW2* within a 1-cM interval between *D12S347* and *D12S368* (Fig. 1).

A physical map comprising the *ORW2* critical region, defined by the flanking markers *D12S347* and *D12S359*, was established (Fig. 4). By adding the sizes

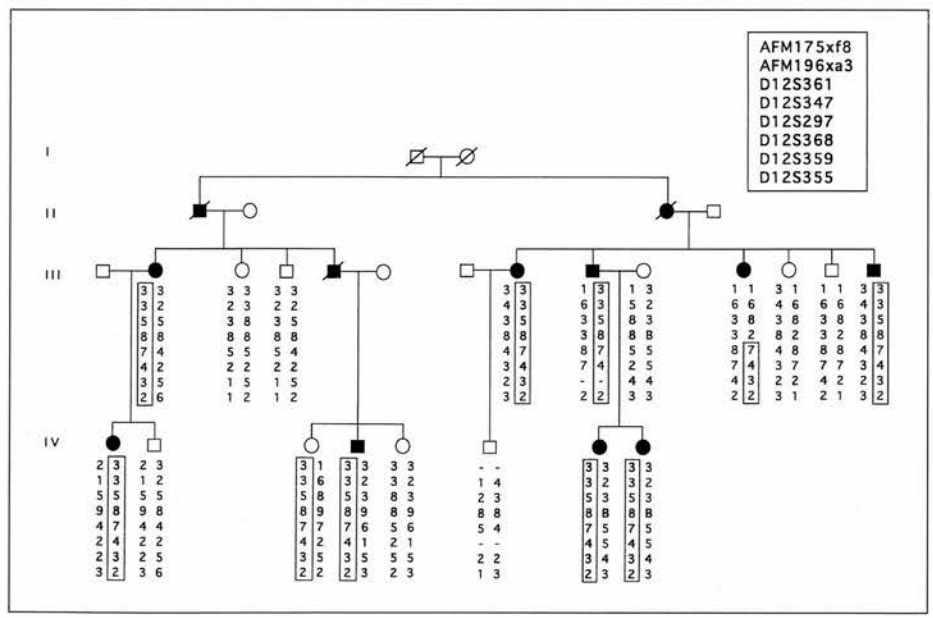


Fig. 2 Haplotype analysis of Family 3 for markers on chromosome 12. The order of markers is shown at the right of the figure. Alleles are number-coded and the disease haplotype is shown boxed. A critical crossover that forms the proximal boundary of the *ORW2* candidate region is shown in generation III, person 11.

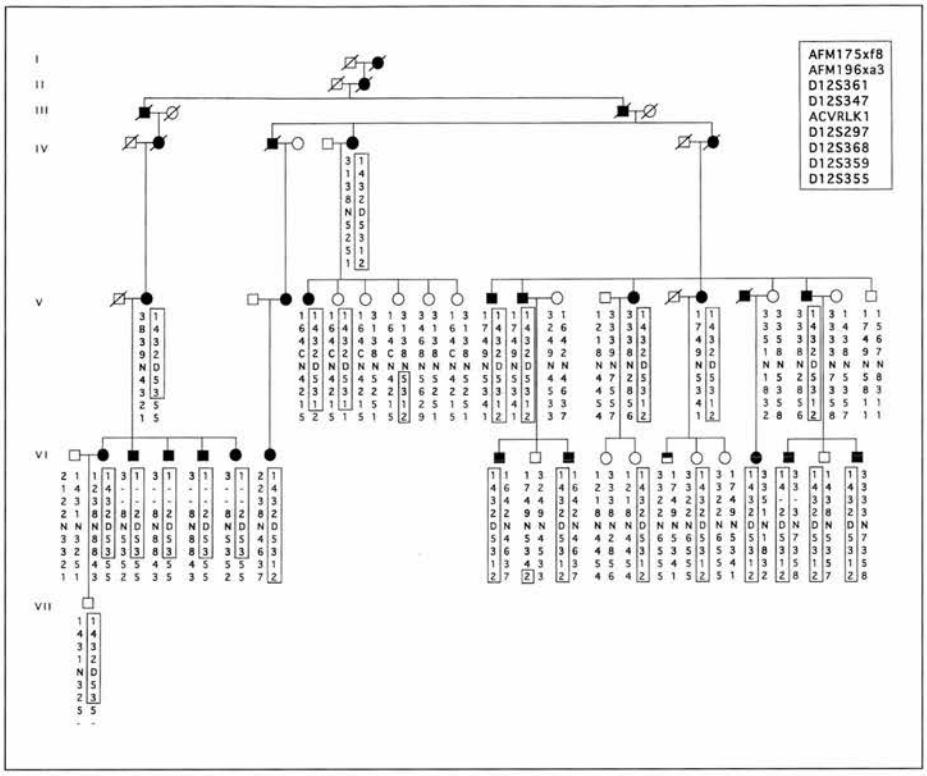


Fig. 3 Haplotype analysis of family 67 for markers on chromosome 12. The order of markers is shown at the right of the figure. Alleles are number-coded and the disease haplotype is shown boxed. A critical crossover that forms the distal boundary of the *ORW2* candidate region is shown in generation V, person 2. All of her affected children show the same recombinant haplotype. The genotype at *ACVRLK1* (*ALK1*) is shown (D, deletion; N, no deletion).

of nonchimaeric YACs that constitute a minimal tiling path between these two markers, the physical size of the critical region is estimated to be 1.38 Mb. A contig of 35 YACs and 2 cosmids in this region contains 36 markers, including *ALK1*, which was mapped previously to 12q11→q14 (ref. 21). *ALK4*, also within the interval, encodes the ubiquitously expressed activin receptor IB[22], and was not considered as strong a candidate for this disorder.

Mutation analysis of *ALK1*

Although genomic sequence for *ALK1* was not available for mutation analysis, two groups have reported sequence for its cDNA[15,23]. Northern blot analysis of different tissues shows two *ALK1* transcripts, of approximately 2.0 and 4.0 kb, expressed in approximately equal proportions. However both research groups have sequenced multiple cDNA clones for *ALK1*, and only a single open reading frame has been found. We obtained cDNA sequence of *ALK1* by RT-PCR of lymphocyte or lymphoblast RNA from four affected individuals, members of families 17, 33, 40, and 67, which show genetic linkage to this region. (Family 3 was unavailable for resampling.)

A 3-bp deletion was detected in the sequence of the amplified cDNA of *ALK1* from an affected member of family 67 (Fig. 5). The deletion segregates with the disease phenotype in this large family (Fig. 3). A few unaffected individuals also have this deletion, but these all showed the disease haplotype throughout the entire candidate interval and are presumed non-penetrant. This deletion in the kinase subdomain II

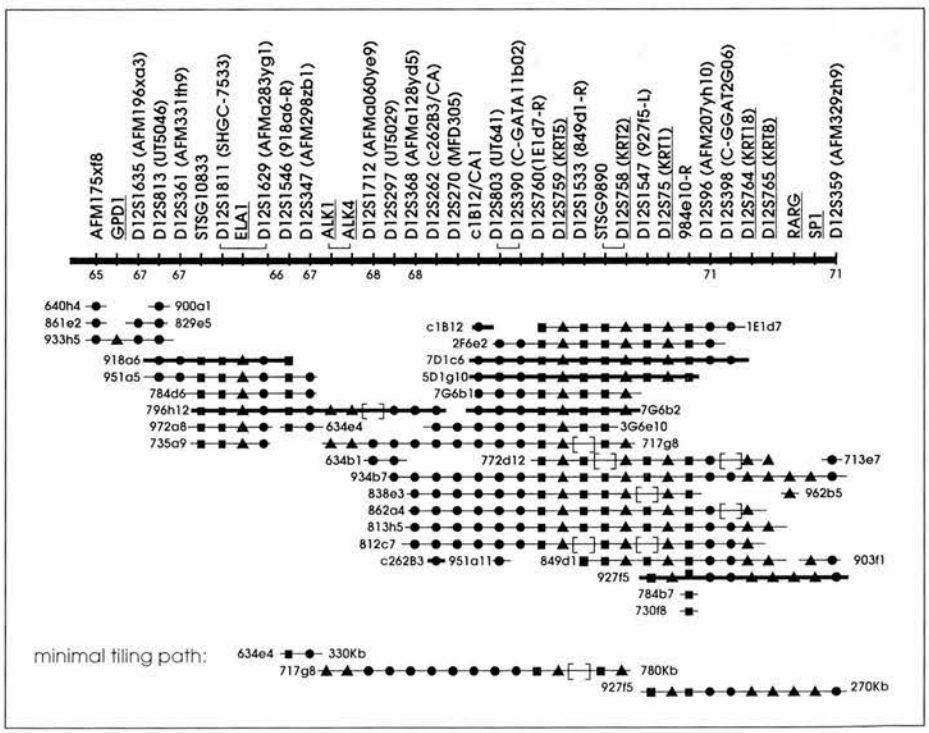


Fig. 4 Physical map of chromosome 12 showing the *ORW2* critical interval and the location of *ALK1*. The horizontal line at the top of the figure represents a portion of the q arm of chromosome 12, with the centromere to the left and qter to the right. Marker names are noted above the horizontal line and correspond to the symbols underneath. Markers whose relative order could not be established unambiguously are enclosed in a bracket. The genetic distance in cM as derived from the Genethon linkage map is indicated for the appropriate markers as the number below the horizontal line. Markers present on individual YACs are shown by corresponding symbols: highly polymorphic markers appear as circles, monomorphic markers are represented as squares, genes are denoted as triangles. YAC addresses are shown next to individual YACs. Square brackets on YACs point to the absence of markers in that position. YACs mapped using FISH are indicated with a thick line.

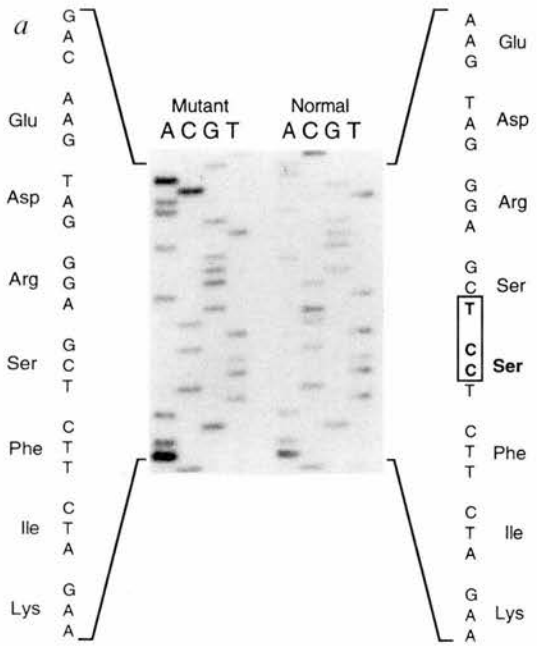


Fig. 5 *a*, cDNA sequence of the middle section of *ALK1* amplified from RNA from the proband of family 67. The original sequencing ladder showed normal sequence up to base number 695 (numbered from the ATG start codon) with a juxtaposition of both normal and deleted sequence beyond this point. In this figure the normal and the mutant sequence are separated (by preferential amplification) and shown side by side. A 3-bp deletion is evident at bases 696 to 698. The deletion causes the loss of a serine residue. *b*, Amino-acid composition of a portion of kinase subdomain II for representative type I serine-threonine kinase receptors[16,24]. Sequences for ALK1-5 (Z22533, Z22534, Z22535, Z22536, and L11695) are from the human[23,31]. ALK6 (Z23143) is from the mouse[17]. ATR1 (activin receptor; U04692), SAX (saxophone; U05209), and TKV (thick veins; U11442) are from *Drosophila*[32–34]. The serine deleted in family 67 is shown in bold-type and highlighted with an asterisk.

(according to the homologous structure protein kinase, PKA-Cα)[24] would result in the loss of one of two adjacent serine residues which are conserved in most of the type I TGF-β receptor family members from mammals to *Drosophila*. The deletion was not found in 390 chromosomes from a panel of unrelated control genomic DNAs.

A point mutation was detected in the amplified cDNA from an affected member of family 33. This base pair substitution segregates with the disease phenotype in the family, and is found in unaffected members only when these individuals show the disease haplotype throughout the candidate region. The mutation was not found in 322 chromosomes in a panel of unrelated control genomic DNAs. The change of a T to G (coding strand) substitutes a positively charged arginine for a highly conserved methionine in the kinase subdomain VIII (Fig. 6).

A point mutation was detected in the amplified cDNA from an affected member of family 40. The mutation co-segregates with the disease haplotype in the family, and was not found in 316 chromosomes in a panel of unrelated control genomic DNAs. The change of a G to A on the coding strand would substitute a

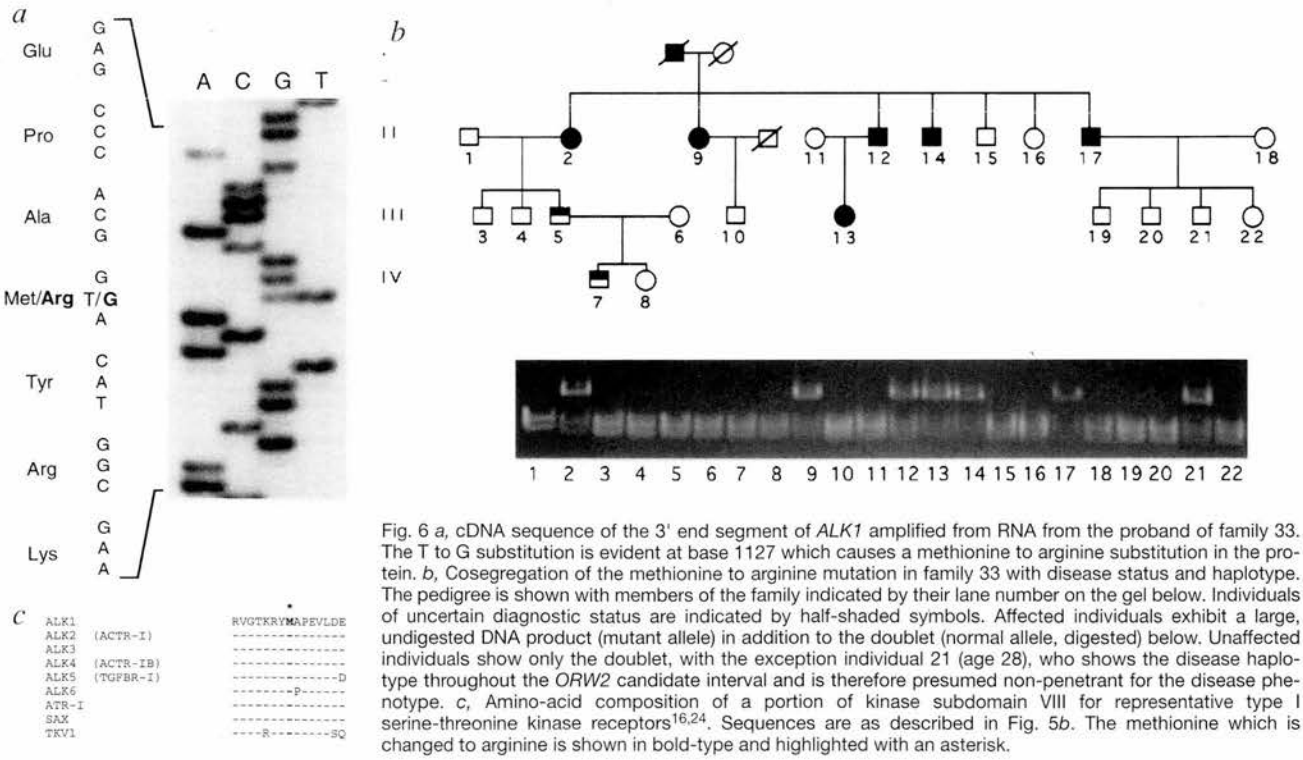


Fig. 6 *a*, cDNA sequence of the 3' end segment of *ALK1* amplified from RNA from the proband of family 33. The T to G substitution is evident at base 1127 which causes a methionine to arginine substitution in the protein. *b*, Cosegregation of the methionine to arginine mutation in family 33 with disease status and haplotype. The pedigree is shown with members of the family indicated by their lane number on the gel below. Individuals of uncertain diagnostic status are indicated by half-shaded symbols. Affected individuals exhibit a large, undigested DNA product (mutant allele) in addition to the doublet (normal allele, digested) below. Unaffected individuals show only the doublet, with the exception individual 21 (age 28), who shows the disease haplotype throughout the *ORW2* candidate interval and is therefore presumed non-penetrant for the disease phenotype. *c*, Amino-acid composition of a portion of kinase subdomain VIII for representative type I serine-threonine kinase receptors[16,24]. Sequences are as described in Fig. 5*b*. The methionine which is changed to arginine is shown in bold-type and highlighted with an asterisk.

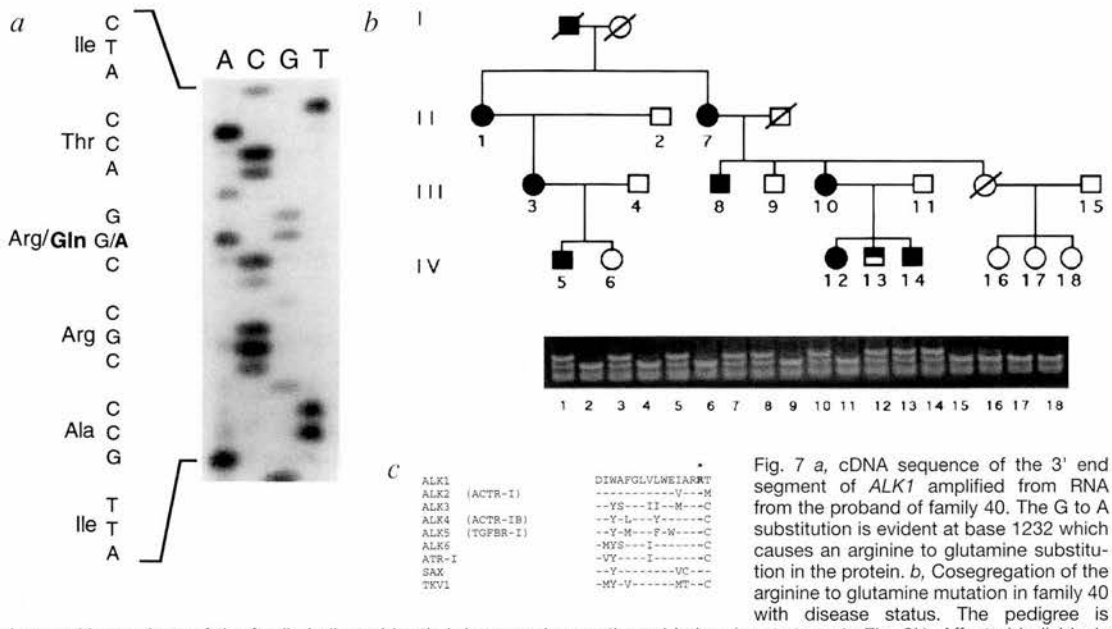


Fig. 7 *a*, cDNA sequence of the 3' end segment of *ALK1* amplified from RNA from the proband of family 40. The G to A substitution is evident at base 1232 which causes an arginine to glutamine substitution in the protein. *b*, Cosegregation of the arginine to glutamine mutation in family 40 with disease status. The pedigree is shown with members of the family indicated by their lane number on the gel below (symbols as in Fig. 6*b*). Affected individuals exhibit a large, undigested DNA product (mutant allele) in addition to the doublet (normal allele, digested) below. Unaffected individuals show only the doublet, with the exception individual 13, who was classified as uncertain diagnostic status. *c*, Amino-acid composition of a portion of kinase subdomain IX for representative type I serine-threonine kinase receptors[16,24]. Sequences are as described in Fig. 5*b*. The arginine which is changed to glutamine in family 40 is shown in bold-type and highlighted with an asterisk.

neutral glutamine for a highly conserved positively charged arginine in kinase subdomain XI (Fig. 7).

Discussion

Reassessment of the ORW2 candidate region defined a maximum 4-cM interval which includes the *ALK1* gene. Three mutations were found in the sequence analysis of *ALK1* in four ORW2 families. All three were identified at the cDNA level, suggesting that these result in functional changes to the protein. We were unable to find the mutation in the RT-PCR products of *ALK1* from the proband of one family. This may be due to a number of factors related to either the stability of the mutant message or the position of the mutation within the *ALK1* gene structure. Alternatively, the mutation in this family may be in a closely linked gene.

Each of the ORW2 mutations might be expected to reduce kinase activity, due to removal or substitution of critical residues in the proposed kinase subdomain structure. The kinase subdomain II of serine-threonine protein kinases is comprised of a highly conserved β-strand 3 and a more variable α-helix B. The invariant lysine (Lys92 in PKA-Cα) in the β-strand 3 is essential for maximal enzyme activity. It interacts with a conserved residue in subdomain III and helps in proper positioning of ATP in the hydrophobic adenine ring pocket by interacting with the α- and β-phosphates. Other residues in subdomain II are also involved in stabilizing ATP, either by direct interaction or by contributing to higher order structure of the kinase necessary for enzyme activity[24]. Thus the deleted serine in the family 67 mutation, in the α-helix B portion of kinase subdomain II, may affect proper binding or positioning of ATP.

The kinase subdomain VIII plays a major role in recognition of peptide substrates, providing a hydrophobic pocket to accommodate hydrophobic residues of typical substrates. The highly conserved alanine-proline-glutamate (A-P-E) motif immediately follows a region of 7–10 residues which are conserved among the members of different protein kinase families[24]. This includes the methionine, normally with M-A-P-E in ALK1 and other

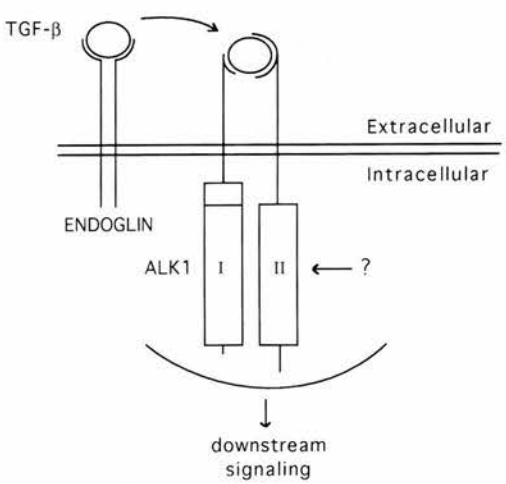


Fig. 8 A model for TGF-β receptor functions involved in ORW (based in part on ref. 35). Endoglin is a homodimeric endothelial cell membrane protein that binds TGF-β. ALK1 is a type I receptor which is also highly expressed in endothelial cells. It can bind TGF-β or activin when co-expressed with the corresponding type II receptor. There is as yet no evidence for an endothelial cell-specific type II TGF-β receptor. In this scenario endoglin sequesters TGF-β and presents the ligand to the signalling complex of ALK1 (a type I) and a type II receptor. Mutations within endoglin and ALK1 have been shown to cause ORW1 and 2 respectively, presumably by disruption of this receptor complex involved in TGF-β signal transduction. Defects in a type II receptor within this complex might also be expected to cause the phenotype of ORW.

type I kinase receptors. This relatively non-polar methionine residue is replaced by a charged arginine residue in the family 33 mutation. This substitution may affect the binding of protein substrate, as the hydrophobicity of the substrate pocket is lessened.

Subdomain IX comprises the large α-helix F, thought to play a role in stabilizing the catalytic loop region, as well as contributing to the hydrophobicity (and thus stability of binding to the protein substrate) of the substrate pocket[24]. The arginine to glutamine substitution in family 40 might affect the stability of this substrate pocket.

The specificity of ligand binding to the receptor complex for the TGF-β superfamily is conferred mainly by the type II receptor. The ALK1 protein can associate with the TGF-β or activin type II receptors after cotransfection in COS cells, with the complex binding TGF-β or activin respectively[15,17,18]. This association suggests that inactivating mutations within the kinase domain of ALK1 might have a dominant-negative effect on TGF-β or activin mediated signalling. Deletion of the kinase domain of the type II TGF-β receptor blocks ligand mediated signalling in a number of systems including renal glomerular capillary endothelial cells[25]. The elucidation of the role of these ORW2 mutations on the receptor complex would require *in vitro* analysis of these mutations on kinase activity, but unfortunately there is no established method to evaluate the kinase activity of ALK1.

The variable response of different cell types to activin and TGF-β may be modulated by tissue-specific expression of multiple receptor isoforms forming the functional signalling complexes. ALK1 and endoglin may interact at the endothelial cell surface in a common signal transduction pathway involving TGF-β (Fig. 8). In this scenario, endoglin (acting in a manner similar to betaglycan[26]) sequesters TGF-β and presents it to an endothelial cell-specific signaling complex containing ALK1. If another ORW locus exists, one candidate might be a novel type II receptor that interacts with ALK1 to form this complex (Fig. 8). Alternatively, ALK1 and endoglin may act in separate pathways (or in ones which converge in the cell) involving different ligands which have similar effects on vascular endothelium.

The vascular lesions in ORW are localized to discreet regions within the affected tissue, with no evidence of abnormal vascular structure or pathology outside the lesions themselves. This suggests that some genetic, physiologic or mechanical event initiates the formation of each vascular lesion. The pathobiology of the disease may be related to remodelling of the vascular endothelium following an unknown initiating event. TGF-β mediates vascular remodelling through effects on extracellular matrix production by endothelial cells, stromal interstitial cells, smooth muscle cells, and pericytes[27]. Responses of endothelial cells to TGF-β depend on interactions with the surrounding extracellular matrix via integrins expressed on their surface, and TGF-β itself regulates integrin expression of endothelial cells[27,28]. Perturbations in the TGF-β signalling pathway in ORW may lead to altered repair of vascular endothelium and remodelling of the vascular tissue via changes in expression profiles of extracellular matrix proteins.

Methods

Clinical evaluation. In total five families were used to generate the data presented in this report. Two of these, families 17 and 33, have been described[9] and are linked to this region of chromosome 12 (ref. 14). Family 67 was also described and is linked to chromosome 12 (ref. 14). However, nine additional members of this family were ascertained for this study, including an individual who contributed a new distal crossover.

Family 3 was described and excludes *ORW1* on chromosome 9 (ref. 10). Our data showed this family is linked to the *ORW2* locus on chromosome 12. In addition, this family provides a critical proximal crossover.

Family 40 is a new family. It is a four generation pedigree for which DNA samples have been collected for seventeen individuals, including seven affected members (Fig. 7b). Epistaxis and muco-cutaneous telangiectases are present in most of the affected individuals. Two of the affected persons reported gastrointestinal bleeding. No affected individuals had any history of pulmonary arteriovenous malformations.

New individuals and families in this study were assessed by medical history and examination of skin, hands, and mucous membranes. A positive diagnosis of ORW was made in the presence of two of the following: affected first degree relative, nosebleeds more than once per month, classical muco-cutaneous lesions, and arteriovenous malformations, with the exception that individuals with only an affected first degree relative and nosebleeds were considered to be of uncertain diagnosis. This approach was suggested[4] as nosebleeds occur frequently in the general population. Individuals were assessed before marker genotype information was available. Peripheral blood (20–30 ml) was obtained with informed consent. DNA extraction was performed as described[14]. RNA was prepared from cultured transformed lymphoblasts or lymphocytes derived from fresh blood, according to protocols furnished with the RNAzol B reagent by the supplier (Tel-Test).

Genotyping and linkage analysis. Analysis of simple sequence repeat markers was performed as described[3]. Two point linkage analysis was also performed as described[9,14].

Physical map construction. A high resolution map of the genetically defined region of *ORW2* was constructed using Yeast Artificial Chromosome (YAC) clones. The map was constructed as described[29]. Several subsets of YAC clones which contained genetically mapped Genethon markers in the *ORW2* interval were initially identified using the Quickmap database. Pooling of these YAC subsets according to their presumed genetic location (Quickmap-based Pooling Strategy) allowed the placement of new, unmapped STSs in less than 60 PCR reactions.

Mutation analysis. The search for mutations in *ALK1* was aided by the availability of the published sequence of the cDNA[15,23]. This allowed specific PCR amplification of cDNA prepared from patient blood lymphocytes or lymphoblasts after reverse transcription with oligo dT. Oligonucleotide primers were designed to amplify the cDNA in three segments, each about 500–750 bp. The 5'-end segment was amplified using a nested approach, with primers 5'–ATTTCCTCTGGGCAGGAGGGAGCC–3' and 5'–AACCTGCCGTGCCACTGTCCTCTG–3' for the first round, and primers 5'–GCGCGTGTCACACTTCATGGCTCTTA–3' and 5'–TGAGCCACTCCCTGTGGTGCAGTC–3' for the second round; the middle segment was amplified with primers 5'–CCTTGCTGGCCCTGGTGGCCCT–3' and CCACTCTCGGGTTGTTGCCGATGTCC–3'; and the 3' segment was amplified using primers 5'–GCGACTTCAAGAGCCGCAAT–3' and 5'–AAGGAATCAGGTGCTCCTGGG–3'. The amplification products were then subjected to direct (cycle) sequencing to detect heterozygotes. The reported cDNA sequence was compared to cDNA sequence derived from one or more affected individuals from each family. When a difference from the reported cDNA

sequence was identified, a corresponding PCR assay was developed to amplify a small segment of genomic DNA containing the putative mutation. These assays were designed to differentially detect the divergent sequences, so that they could be used to determine if the putative mutation cosegregated with the disease locus in the corresponding ORW2 family, and to derive data on frequency of the mutation in random unaffected individuals.

Primers used to amplify the region flanking the deletion mutation in family 67 were 5'–TGTGGCACGGTGAGAGTGTGGCC–3' and 5'–CCCGGAACCAGGACTGTTCATCCCTC–3'. Primers used to amplify the substitution mutation in family 33 were 5'–ACTCACAGGGCAGCGATTACCT–3' and 5'–GTAGGACTCAAAGCAGTCCGTGC–3'; the mutation destroys an *Nla*III site; this assay was used to show cosegregation with disease haplotype in family 33. This mutation also creates a *Bsp*1286I site, allowing the detection of an extra band after restriction of the radiolabelled PCR product (used in the survey of unrelated individual DNAs). Primers used to amplify genomic DNA containing the mutation in family 40 were 5'–CACGGACTGCTTTGAGTCCTACAAGT–3' and 5'–CACGATGCCATTCACGATGGT–3'; the mutation destroys an *Hpa*II site, allowing detection of a larger uncut fragment in the digest of affected members of family 40.

Control DNA samples were obtained from families seeking genetic testing for either haemophilia or Duchenne muscular dystrophy as described[30].

Acknowledgments

We gratefully acknowledge the family members who participated in this study, as well as the enthusiastic support of the Hereditary Hemorrhagic Telangiectasia Foundation, International. We thank M. R. Tudor and L. M. Beaty for help in family ascertainment, M. Qumsiyeh for cytogenetic analysis, and L. Brody and C. Boehm for control population DNA samples. This study was supported by NIH grant HL 49171, a Baxter Foundation award, and a Grant-in-Aid (95006970) and Established Investigator Award (95002940) from the American Heart Association to D.A.M. Support for D.W.J. and T.T.S. comes from NIH fellowship awards 1F32 HL09349 and 1F32 HL09394, respectively. R.K. is supported by NIH (HG00965) and a Cancer Center Grant to A.E.C.O.M.

Received 21 March; accepted 29 April 1996.

1. Guttmacher, A.E., McKinnon, W.C. & Upton, M.D. Hereditary hemorrhagic telangiectasia: a disorder in search of the genetics community. *Am. J. Med. Genet.* **52**, 252–253 (1994).
2. Guttmacher, A.E., Marchuk, D.A. & White, R.I. Current concepts: Hereditary hemorrhagic telangiectasia. *New Engl. J. Med.* **333**, 918–924 (1995).
3. McDonald, M.T. *et al.* A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33–34. *Nature Genet.* **6**, 197–204 (1994).
4. Shovlin, C.L. *et al.* A gene for hereditary haemorrhagic telangiectasia maps to chromosome 9q3. *Nature Genet.* **6**, 205–209 (1994).
5. McAllister, K.A. *et al.* Endoglin, a TGF-β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nature Genet.* **8**, 345–351 (1994).
6. Gougos, A. & Letarte, M. Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J. Biol. Chem.* **265**, 8361–8364 (1990).
7. Cheifetz, S. *et al.* Endoglin is a component of the transforming growth factor-β receptor system in human endothelial cells. *J. Biol. Chem.* **267**, 19027–19030 (1992).
8. McAllister, K.A. *et al.* Six novel mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type 1 suggest a dominant-negative effect of receptor function. *Hum. Mol. Genet.* **4**, 1983–1985 (1995).
9. McAllister, K.A. *et al.* Genetic heterogeneity in hereditary hemorrhagic telangiectasia: possible correlation with clinical phenotype. *J. Med. Genet.* **31**, 927–932 (1994).
10. Porteous, M.E.M. *et al.* Genetic heterogeneity in hereditary haemorrhagic telangiectasia. *J. Med. Genet.* **31**, 925–926 (1994).
11. Heutink, P. *et al.* Linkage of hereditary hemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. *J. Med. Genet.* **31**, 933–936 (1994).
12. Berg, J.N., Guttmacher, A.E., Marchuk, D.A. & Porteous, M.E.M. Clinical heterogeneity in hereditary haemorrhagic telangiectasia: are pulmonary arteriovenous malformations more common in families linked to endoglin? *J. Med. Genet.* **33**, 256–257 (1996).
13. Vincent, P. *et al.* A third locus for hereditary haemorrhagic telangiectasia maps to chromosome 12q. *Hum. Mol. Genet.* **4**, 945–949 (1995).
14. Johnson, D.W. *et al.* A second locus for hereditary hemorrhagic telangiectasia maps to chromosome 12. *Genome Res.* **5**, 21–28 (1995).
15. Attisano, L. *et al.* Identification of human activin and TGF-β type I receptors that form heteromeric kinase complexes with type II receptors. *Cell* **75**, 671–680 (1993).
16. ten Dijke, P. *et al.* Serine/threonine kinase receptors. *Prog. Growth Factor Res.* **5**, 55–72 (1994).
17. ten Dijke, P. *et al.* Characterization of type I receptors for transforming growth factor-b and activin. *Science* **264**, 101–104 (1994).
18. Miyazono, K., ten Dijke, P., Yamashita, H. & Heldin, C.-H. Signal transduction via serine/threonine kinase receptors. *Semin. Cell Bio.* **5**, 389–398 (1994).
19. Mathews, L.S. Activin receptors and cellular signaling by the receptor serine kinase family. *Endocrine Rev.* **15**, 310–325 (1994).
20. Gyapay, G. *et al.* The 1993–94 Genethon human genetic linkage map. *Nature Genet.* **7**, 246–339 (1994).
21. Kucherlapati, R., Craig, I. & Marynen, P. Report of the second international workshop on human chromosome 12 mapping 1994. *Cytogenet. Cell Genet.* **67**, 246–264 (1994).
22. Carcamo, J. *et al.* Type I receptors specify growth-inhibitory and transcriptional responses to transforming growth factor β and activin. *Mol. Cell. Biol.* **14**, 3810–3821 (1994).
23. ten Dijke, P. *et al.* Activin receptor-like kinases: a novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. *Oncogene* **8**, 2879–2887 (1993).
24. Hanks, S.K. & Hunter, T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576–596 (1995).
25. Choi, M.E. & Ballermann, B.J. Inhibition of capillary morphogenesis and associated apoptosis by dominant negative mutant transforming growth factor-β receptors. *J. Biol. Chem.* **270**, 21144–21150 (1995).
26. Lopez-Casillas, F., Wrana, J.L., & Massague, J. Betaglycan presents ligand to the TGF-β signaling receptor. *Cell* **73**, 1435–1444 (1993).
27. Madri, J.A. *et al.* Interactions of matrix components and soluble factors in vascular responses to injury. Modulation of cell phenotype. In *Endothelial cell dysfunctions* (eds Simionescu, N. & Simionescu, M.) (Plenum Press, New York, 1992).
28. Luscinkas, F.W. & Lawler, J. Integrins as dynamics regulators of vascular function. *FASEB J.* **8**, 929–938 (1994).
29. Krauter, K. *et al.* A second-generation YAC contig map of human chromosome 12. *Nature* **377**, 321–333 (1995).
30. Struewing, J.P. *et al.* The carrier frequency of the BRCA1 185delAG mutation is approximately 1 percent in Ashkenazi Jewish individuals. *Nature Genet.* **11**, 198–200 (1995).
31. Franzen, P. *et al.* Cloning of a TGF-β type I receptor that forms a heteromeric complex with the TGF-β type II receptor. *Cell* **75**, 681–692 (1993).
32. Wrana, J.L. *et al.* Two distinct transmembrane serine/threonine kinases from Drosophila melanogaster form an activin receptor complex. *Mol. Cell. Biol.* **14**, 944–950 (1994).
33. Xie, T., Finelli, A.L. & Padgett, R.W. The Drosophila saxophone gene encodes a serine-threonine kinase receptor of the TGF-β superfamily. *Science* **263**, 1756–1759 (1994).
34. Brummel, T.J. *et al.* Characterization and relationship of Dpp receptors encoded by the saxophone and thick veins genes in Drosophila. *Cell* **78**, 251–261 (1994).
35. Wrana, J.L. *et al.* Mechanism of activation of the TGF-β receptor. *Nature* **370**, 341–347 (1994).

The Activin Receptor–Like Kinase 1 Gene: Genomic Structure and Mutations in Hereditary Hemorrhagic Telangiectasia Type 2

Jonathan N. Berg,[1,3] Carol J. Gallione,[1] Timothy T. Stenzel,[1,2] David W. Johnson,[1] William P. Allen,[4] Charles E. Schwartz,[4] Charles E. Jackson,[5] Mary E. M. Porteous,[3] and Douglas A. Marchuk[1]

Departments of [1]Genetics and [2]Pathology, Duke University Medical Center, Durham; [3]Department of Human Genetics, University of Edinburgh, Edinburgh; [4]Greenwood Genetics Center, Greenwood, SC; and [5]Department of Medicine, Division of Clinical and Molecular Genetics, Henry Ford Hospital, Detroit

Summary

The activin receptor–like kinase 1 gene (ALK-1) is the second locus for the autosomal dominant vascular disease hereditary hemorrhagic telangiectasia (HHT). In this paper we present the genomic structure of the ALK-1 gene, a type I serine-threonine kinase receptor expressed predominantly in endothelial cells. The coding region is contained within nine exons, spanning <15 kb of genomic DNA. All introns follow the GT-AG rule, except for intron 6, which has a TAG|gcaag 5′ splice junction. The positions of introns in the intracellular domain are almost identical to those of the mouse serine-threonine kinase receptor TSK-7L. By sequencing ALK-1 from genomic DNA, mutations were found in six of six families with HHT either shown to link to chromosome 12q13 or in which linkage of HHT to chromosome 9q33 had been excluded. Mutations were also found in three of six patients from families in which available linkage data were insufficient to allow certainty with regard to the locus involved. The high rate of detection of mutations by genomic sequencing of ALK-1 suggests that this will be a useful diagnostic test for HHT2, particularly where preliminary linkage to chromosome 12q13 can be established. In two cases in which premature termination codons were found in genomic DNA, the mutant mRNA was either not present or present at barely detectable levels. These data suggest that mutations in ALK-1 are functionally null alleles.

Introduction

Hereditary hemorrhagic telangiectasia (HHT) is an autosomal dominant disease of blood vessels that is characterized by severe recurrent nosebleeds, mucocutaneous telangiectases, gastrointestinal hemorrhage, and a high incidence of vascular malformations in the lung and brain (Guttmacher et al. 1995). Most of the published families with HHT fall into two linkage groups: HHT1, mapping to chromosome 9q33 (McDonald et al. 1994; Shovlin et al. 1994); and HHT2, mapping to chromosome 12q13 (Johnson et al. 1995; Vincent et al. 1995). Patients with HHT linked to 9q33 may have a higher incidence of symptomatic pulmonary arteriovenous malformations than is seen in those with HHT linked to 12q13 (Heutink et al. 1994; McAllister et al. 1994*b*; Porteous et al. 1994; Berg et al. 1996), suggesting that, in the pulmonary vasculature, the HHT1 gene has an additional role not shared by HHT2. A third rare variant of HHT has been reported in one large family with hepatic involvement as the major manifestation (Piantanida et al. 1996), with exclusion of linkage to both chromosome 9 and chromosome 12.

The genes involved at the two major loci have been identified. HHT1 is the endoglin gene on chromosome 9 (McAllister et al. 1994*a*); HHT2 is the ALK-1 gene on chromosome 12 (Johnson et al. 1996). Both of these genes are members of the TGF-β receptor superfamily (Cheifetz et al. 1992; Attisano et al. 1993; ten Dijke et al. 1993). The ALK-1 protein has the properties of a type I serine-threonine kinase receptor (Attisano et al. 1993; ten Dijke et al. 1993). It has been shown to bind either activin or TGF-β in the presence of their respective type II receptors but does not bind ligand alone (Attisano et al. 1993). The mechanism for downstream signaling of ALK-1 has yet to be elucidated, but it seems probable that it involves pathways similar to those implicated in the signaling of other serine-threonine kinase receptors (Niehrs 1996; Zhang et al. 1996). Expression studies of cell lines by reverse-transcriptase–PCR (RT-PCR) have shown that significant transcription of ALK-1 occurs only in endothelial cell lines (Attisano et al. 1993), although we have demonstrated that it is possible to amplify ALK-1 message from peripheral blood leukocyte mRNA. This has permitted identification of three

Received November 20, 1996; accepted for publication April 25, 1997.

Address for correspondence and reprints: Dr. Douglas A. Marchuk, Department of Genetics, Duke University Medical Center, Box 3175, Durham, NC 27710. E-mail: march004@mc.duke.edu

0002-9297/97/6101-0011$02.00

mutations in ALK-1 in unrelated individuals affected with HHT2 (Johnson et al. 1996).

In this paper we have identified the size and position of introns in the coding portion of the ALK-1 gene. We have developed PCR assays to amplify each exon of the coding region of ALK-1, from genomic DNA. The ALK-1 gene was sequenced in a panel of 12 unrelated patients with HHT, 6 of whom have either evidence of HHT linkage to markers on 12q13 or exclusion of linkage of HHT to 9q34; for the remaining 6 insufficient linkage data were available. The position and type of each mutation found are discussed in relationship to both the function of the protein and the pathobiology of the disease.

Subjects, Material, and Methods

Human Subjects

Individuals with HHT had been assessed in, and DNA samples had obtained from, families 1–5, 17, and 92, as described elsewhere (McAllister et al. 1994*b*; Porteous et al. 1994; Johnson et al. 1995; Berg et al. 1996), after informed consent had been obtained. Single samples had also been collected from four unrelated individuals by use of the same methods.

Identification of a Genomic P1 Artificial Chromosome (PAC) Clone Containing ALK-1

DNA from pools of a two-dimensional arrayed PAC library (Ioannou et al. 1994) was screened by use of the primers 5′-CGCGTGTCACACTTCATGGCTC-3′ and 5′-ATCAGAAGGCCTTTCCTGGGGG-3′, designed to the 5′ UTR of ALK-1 published by Attisano et al. (1993. The positive clone from a single positive 384-well plate (plate 265) was identified at position C14 by hybridization using standard protocols (Sambrook et al. 1989). This clone was designated "265C14." Subclones were generated by use of *Bgl*II-digested PAC cloned into *Bam*HI-digested pBluescript phagemid (Stratagene).

Localization of Introns

With a range of primers designed from the published cDNA sequence, PCR reactions were performed to amplify segments of genomic DNA containing the ALK-1 gene. Whenever a PCR fragment generated was larger than that predicted from the cDNA sequence, it was assumed to contain an intron and was sequenced. Where sequencing was unclear, further sequence was generated from the plasmid subclones by use of different primers designed for the ALK-1 cDNA sequence. Intron size was estimated by size of PCR product, or, for three of the introns, sequence was generated for the entire intron.

PCR Amplification of Individual Exons

With the intronic sequence generated, primer pairs were designed to amplify each exon of the coding part of the ALK-1 gene. Primer sequence and conditions are given in table 1. PCR was carried out for 35 cycles: 95°C for 45 s, annealing temperature for 45 s, and 72°C for 1 min. One hundred nanograms of each primer, 1 × buffer, 50 ng of genomic DNA, 0.2 mM dNTPs, and 1 unit of Gibco-BRL *Taq* polymerase were used in each 25-μl reaction. The buffer used was either that supplied by Boehringer Mannheim Biochemicals (BMB) or that originally used by Weissenbach (W; Gyapay et al. 1994); in both cases the magnesium ion concentration was 1.5 mM.

Mutation Analysis

For each patient in the study, each exon was amplified and sequenced. Sequencing was performed by use of the Amersham Thermosequenase cycle sequencing kit, which uses P^{33}-labeled dideoxy terminators to reduce artifactual bands. Samples of each nucleotide-specific termination reaction were loaded in adjacent lanes, making a mutation in any single lane very easy to identify. In one instance, for family 92, it was necessary to clone the exon into PCR Script by using the Invitrogen TA cloning kit, prior to sequencing. Conservation of the region affected by the mutation was determined by use of the BLAST algorithm, to match the ALK-1 region containing the mutation against the Genbank nonredundant protein database.

RT-PCR

For RT-PCR, peripheral blood leukocyte mRNA was reverse transcribed by use of a mixture of oligo dT and random hexamers. A 250-bp segment of ALK-1 exon VII was amplified by use of primers 5′-TGTGGCACG-GTGAGAGTGTGGCC-3′ and 5′-GTTGCTCTTGAC-CAGCACAT-3′. A minus-reverse-transcriptase control was included.

Polymorphism Screen

For each mutation, a panel of 100–120 normal individuals was screened to ensure that the mutation did not arise as a population polymorphism. The screening was performed by altered restriction-enzyme site, where possible. In one case in which no enzyme site was altered, the mutant PCR product gave a heteroduplex band that could be resolved by gel electrophoresis as described elsewhere (McAllister et al. 1995). For two mutations it was necessary to screen the population samples by sequencing the PCR product. In these instances, only a single termination reaction was run for each individual screened.

Results

Genomic Structure of ALK-1

A single PAC clone containing the ALK-1 gene was isolated from a human genomic PAC library (Ioannou

Table 1

Primers for Amplification of ALK-1 Gene Exons

Exon	5′ Assay Primer	3′ Assay Primer	Annealing Temperature (°C)	Buffer	Size (bp)
II	CTCTGTGATTTCCTCTGGGCA	TACATTCTCCCCAGCTTCTCAA	62[a]	BMB	266
III	AGCTGGGACCACAGTGGCTGA	GGAGGCAGGGGCCAAGAAGAT	64	W	345
IV	AGCTGACCTAGTGGAAGCTGA	CTGATTCTGCAGTTCCTATCTG	60	W	318
V	AGGAGCTTGCAGTGACCCAGCA	ATGAGAGCCCTTGGTCCTCATCCA	68[a]	BMB	242
VI	AGGCAGCGCAGCATCAAGAT	AAACTTGAGCCCTGAGTGCAG	60	BMB	294
VII	TGACGACTCCAGCCTCCCTTAG	CAAGCTCCGCCCACCTGTGAA	65[a]	W	388
VIII	AGGTTTGGGAGAGGGGCAGGAGT	GGCTCCACAGGCTGATTCCCCTT	65	BMB	293
IX	TCCTCTGGGTGGTATTGGGCCTC	CAGAAATCCCAGCCGTGAGCCAC	68[a]	BMB	256
X	TCTCCTCTGCACCTCTCTCCCAA	CTGCAGGCAGAAAGGAATCAGGTGCT	65[a]	BMB	197

[a] Hot start improves amplification.

et al. 1994). This PAC clone was designated "265C14." In Southern blot analysis, cDNA from the entire coding region of ALK-1 hybridized to genomic fragments totaling 15 kb. The position of each intron and the size of each exon, as determined by PCR amplification from the PAC clone, are shown in figure 1. Exon I contains the 5′ untranslated sequence published by ten Dijke et al. (1993) and was shown, by sequencing of a subclone, to be present in 265C14.

The 5′ untranslated sequence of ALK-1 differs in the two published descriptions of this gene (Attisano et al. 1993; ten Dijke et al. 1993). Genomic sequencing shows that the 5′ untranslated sequence published by Attisano is part of exon II. ten Dijke et al.'s version is a splice variant arising from splicing of exon I to a consensus splice junction present 7 bp upstream of the start codon in exon II, where identity with Attisano et al.'s version of this sequence begins. This splice is marked with an arrow in figure 1. A consensus splice is present at the 3′ side of exon I. Both splice variants can be amplified by RT-PCR of mRNA extracted either from endothelial cell lines or from peripheral blood leukocytes.

The coding region of ALK-1 is contained within nine exons. All except one intron follow the GT-AG rule. Intron 6, between exons VI and VII, has a nonconsensus sequence at the 5′ junction. The most likely position for the splice, however, gives the splice-junction sequence TAG|gcaag, which is the sequence most commonly reported for nonimmunoglobulin eukaryotic nonconforming 5′ splice sites (Senapathy et al. 1990). The 3′ splice of this intron would then have the usual consensus sequence, cag|G. Sequence data from all the intron-exon borders of the coding region are available as Genbank accession numbers U77707–U77713.

Comparison with a Murine Type I Kinase

Another type I serine-threonine kinase receptor for which the genomic structure has been published is the mouse receptor TSK-7L (Ebner et al. 1993; Schmitt et al. 1995), a homologue of human ALK-2 and activin type I receptor. The amino acid sequence of the kinase domain is 79% identical to human ALK-1, but the extracellular domain is only 23% identical. Comparison of the intron-exon positions of ALK-1 with those of TSK-

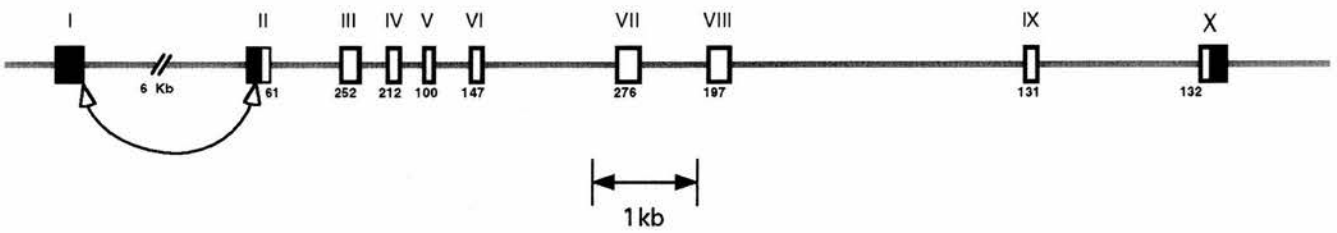


Figure 1 Diagrammatic representation of the ALK-1 gene, approximately to scale. The exons are marked by boxes—unblackened for the coding regions and black for the noncoding regions. The exon number referred to in the text is given as a roman numeral above the exon. Exact exon size (in bp) is given below the exon, in arabic numerals. The curved arrow marks the putative alternative splice, in the 5′ UTR, that would lead to the 5′ sequence published by ten Dijke et al. (1993). For the 5′ Attisano et al. (1993) sequence, transcription would start farther 5′ in exon II.

7L shows almost exact conservation of the position of introns within the kinase domain. It is notable that four of the intron-exon splices occur between the first and second bases of a codon in the human sequence. Wherever this happens, the amino acid involved is uncharged. This is also the case with the murine receptor. Although the location of intron 6 is exactly conserved between mouse and human, the mouse intron follows the GT-AG rule, whereas the human intron 6 has a nonconsensus sequence, as outlined above.

Novel Mutations Detected in the ALK-1 Gene

The entire coding portion and intron-exon borders of the ALK-1 gene were sequenced in 12 unrelated individuals affected with HHT. Mutations were found in ALK-1 in nine of the patients. Of these, six were from families showing either strong evidence of HHT linkage to chromosome 12 (either a LOD score >3.0, with closely linked markers, or a LOD score close to the maximum predicted [in the previously published families 2, 3, and 17] or exclusion of chromosome 9 linkage with inconclusive evidence of linkage to chromosome 12 [in previously published families 4 and 5 and new family 92]). Mutations were also found in three of six patients with HHT who were from small families for which there were inclusive linkage data. For each mutation, 100–120 unrelated individuals were analyzed, to show that the sequence change was a true mutation and not a population polymorphism. When other family members were available, the mutation was shown to be present in the other affected individuals from the same kindred. The mutations and their positions in the gene are described in table 2. Figure 2 shows sequence data and cosegregation of the ALK-1 mutation with disease status in a part of the previously unpublished family 92.

For two of the three individuals for whom no mutation was found, a genomic Southern blot of DNA digested with *Hin*dIII, *Eco*RI, or *Bgl*II and probed with ALK-1 cDNA showed no evidence of a gross genomic rearrangement. DNA available from the third individual was insufficient for Southern blot analysis.

All 12 mutations seen so far (9 described here and 3 described elsewhere) leave the putative signal peptide and the transmembrane domain of ALK-1 intact. Two mutations lead to the change of a conserved amino acid in the extracellular domain. Two mutations would lead to premature truncation of the ALK-1 protein just after the transmembrane domain, if translated. The remaining eight mutations either change residues conserved in the serine-threonine kinase family or lead to a frameshift and premature stop in the kinase domain.

In order to determine whether a mutant protein is required for the pathogenesis of the disease, we investigated the presence of mutant transcripts in patient mRNA. In two cases in which a premature termination codon is created in exon VII by either a frameshift or a nonsense mutation, mRNA was isolated from peripheral blood leukocytes from an affected family member and was amplified by RT-PCR, by use of a single round of PCR. In both cases, only the wild-type allele was represented at any appreciable level in the product. These procedures were replicated, with the same result. Figure 3 shows this for a frameshift mutation in exon VII, where no mutant transcript is visible.

Discussion

The position of introns within the kinase domain is very highly conserved between mouse Tsk7L and human ALK-1. In both genes, four of the splice sites fall between the first and second bases of a codon that encodes a neutral amino acid. It has been hypothesized that this allows alternative splicing of the kinase domain, which are still in frame but have variations in kinase activity (Schmitt et al. 1995). Intron 6 of the ALK-1 gene has a variant splice sequence at the 5′ side—AG/gcaag, instead of the usual AG/gt. This has been observed in other genes (Senapathy et al. 1990). The 3′ splice junction is the usual consensus, cag/G. The function of this variant splice sequence is unknown, but it may represent a means for controlling the amount of ALK-1 mRNA that is spliced. This splice-sequence variation is not seen in the mouse, but the splice location occurring between the first and second bases of a glycine codon is the same.

Mutations were found in ALK-1 in all three of the families in which the disease is known to be linked to the ALK-1 region, as well as in all three families in which linkage to HHT1 on chromosome 9 had been excluded. When no linkage data were available, mutations were found in three of the six individuals analyzed. The three patients in whom no mutation was found may come from families in which the disease is linked to other loci, or they may have a form of mutation not easily detected by PCR and sequencing of the coding exons.

Genomic sequencing may prove a more reliable diagnostic test than RT-PCR from peripheral blood leukocyte mRNA. A low level of the ALK-1 transcript in leukocytes in addition to unstable mutant mRNA of some mutations may lead to difficulties in detection of the mutation by RT-PCR. Our comparatively high rate of mutation detection and the small size of the ALK-1 gene make genomic sequencing a viable diagnostic test for HHT2, particularly if a family has previously shown linkage of the disease to chromosome 12q13. Since a clinical diagnosis of HHT in young adults is often uncertain, presymptomatic diagnosis of HHT will allow screening for the serious complications to be targeted only to those at risk.

Table 2

Summary of ALK-1 Mutations in HHT2

Exon	Protein Region	Mutation Found (Position)	Effect on Protein
II	Start	None	
	Signal peptide		
III	Extracellular domain	G→T (150)	Trp (conserved)→Cys
		G→A (200)	Arg (conserved)→Gln
IV	Extracellular domain	G→A (423)	Premature STOP (intracellular)
	Transmembrane domain	G→T (475)	Premature STOP (intracellular)
	Start of intracellular domain		
V	Glycine-Serine domain	None	
VI	ATP binding site	3-bp deletion (694)[a]	Deletion Ser in frame
	Start of kinase domain		
VII	Kinase domain	Ins T (865)	Frame shift and premature STOP
		C→A (924)	Premature STOP
		G→T (998)	Ser (conserved)→Ile
VIII	Kinase domain	C→T (1120)	Arg (conserved)→Trp
		T→G (1126)[a]	Met (conserved)→Arg
		G→A (1232)[a]	Arg (conserved)→Gln
IX	Kinase domain	C→A (1270)	Pro (conserved)→Thr
X	End of kinase domain	None	

[a] Previously published by Johnson et al. (1996).

All 12 mutations identified in ALK-1 leave the signal peptide and transmembrane domain intact. Two extracellular mutations cause changes of conserved amino acids, and there are two premature stops before the glycine-serine domain. The remaining mutations disrupt the kinase domain either by amino acid change or by small insertion/deletion. Such mutations may function in one of several ways, including a dominant-negative effect of the mutant ALK-1 receptor, haploinsufficiency of ALK-1 protein, or complete loss of functional ALK-1 because of a second somatic mutation in the normal allele.

The hypothesis of a dominant-negative effect requires that an altered protein be transcribed and translated from the mutant allele—and that this protein interfere with the function of the heteromeric receptor signaling complex. Support for this mechanism includes the conservation of signal peptide and transmembrane regions in all mutations found to date. If translated, the abnormal ALK-1 protein may retain the ability to bind a type II receptor in the presence of ligand; but it may not be able to transduce signal, either because of disruption of the kinase domain or because of failure to bind ligand because of alteration of the extracellular domain. Similar mutations in another type I receptor show a dominant-negative effect. Wild-type Mv1Lu cells, which express the type I TGF-β receptor, show reduced TGF-β signaling after transfection with an inactivated TGF-β type I receptor (Feng et al. 1995; Yamamoto et al. 1996), created either by deletion of the cytoplasmic domain or by point mutation in the kinase domain. This dominant-negative effect is even more pronounced with truncated forms of the type II receptor.

However, RT-PCR of two premature protein-truncating mutations observed in exon VII exhibit very low to undetectable levels of the mutant transcript from cDNA of peripheral blood leukocytes, suggesting that some mutant transcripts are unstable. Vascular endothelial cells from patients were unavailable for this expression study. Nonetheless, these data suggest that a mutant protein product may not be present for at least some ALK-1 mutations and that a dominant-negative effect is not required for the pathogenesis of HHT.

The vascular lesions observed in HHT may develop in response to reduced levels of ALK-1 that are due to haploinsufficiency. The mutations identified to date are consistent with the hypothesis that they are all functionally null alleles. The inheritance of a single mutant copy would then predispose the individual to develop the vascular lesions observed in HHT. Lesion formation itself could be due to local environmental or mechanical effects at the lumen of the vessel or to additional genetic alterations.

Finally, it is possible that the development of a vascular lesion initiates with a complete loss of ALK-1 signaling in an endothelial cell. This could be due to somatic mutation of the normal ALK-1 allele. Some support for the possibility of a "second hit" is given by the apparently random distribution of discrete lesions, the number of which increases with age. Mucocutaneous telangiec-

tases appear abruptly and are progressive (Plauchu et al. 1989). However, one study has shown that the classical telangiectases arise as a result of dilatation of blood vessels (Braverman et al. 1990) and do not appear to be proliferative lesions. Also, studies on the development and growth of pulmonary arteriovenous malformations have shown gradual size increases with time, requiring as long as a decade or more to double in diameter (Vase et al. 1985; Masakazu et al. 1988). These studies suggest that, if ALK-1 requires a somatic second mutation for the formation of a vascular lesion, it is not acting as a tumor-suppressor gene. Since TGF-β signaling in endothelial cells modulates vascular remodeling by inducing changes in the extracellular matrix (Madri et al. 1989; Merwin et al. 1990), complete loss of ALK-1 signaling may induce remodeling of the vascular bed, rather than directly affect the rate of endothelial cell proliferation. The proof of this mechanism would require detection of either loss of heterozygosity or somatic mutation of the normal ALK-1 allele in the single layer of endothelium removed from an HHT-associated vascular lesion. In a similar analysis of the renal cysts seen in autosomal dominant polycystic kidney disease type 1, loss of heterozygosity and somatic mutation of the normal PKD1 allele were observed only when methods were developed to isolate the single layer of epithelium lining the renal cysts (Qian et al. 1996).

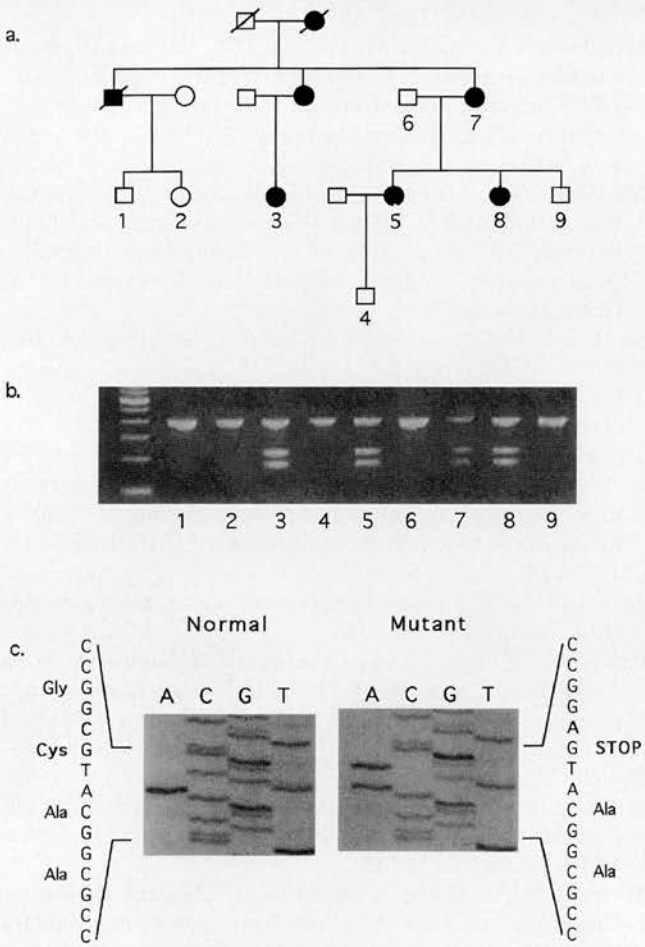


Figure 2 Molecular analysis of the ALK-1 mutation in family 92. *a*, Portion showing affected family members. The disease status cosegregates with presence of a *StuI* site in the exon VII product shown in panel *b*. *b*, *StuI* digest of the exon VII PCR product from family 92. The mutation creates a site that leads to generation of fragments of ~170 and ~210 bp, as well as to an undigested fragment of 380 bp. In unaffected individuals, only the undigested fragment is seen. A 100-bp ladder is used as the size standard. *c*, Sequence from an affected member of family 92. Exon VII product was subcloned as described and was sequenced. Sequences from mutant and wild-type clones are shown. In the mutant clone a C→A base change causes a premature stop. This base change creates a *StuI* site in the mutant exon VII product.

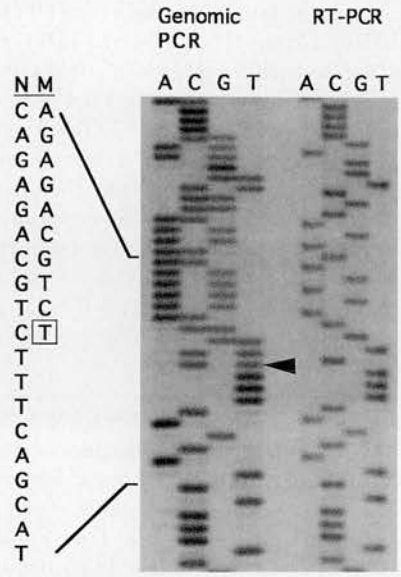


Figure 3 Unstable transcript of the mutant ALK-1 allele in family 17. DNA and mRNA from peripheral blood leukocytes of an affected individual (family 17) were used in genomic PCR and RT-PCR, respectively. Both PCR products were then sequenced by use of an identical primer within the exon. The sequence of the normal (N) allele is apparent in both samples, whereas that of the mutant (M) allele is seen only in the genomic PCR product. Insertion of a T (boxed in the sequence at the left, indicated by an arrow to the right of the band) produces, beyond the inserted nucleotide, superimposed sequences of the normal and mutant alleles, only in the genomic product.

Acknowledgments

We are grateful to Dr. P. de Jong for a copy of the PAC library, to Mary Robyn Tudor for assistance in preparing the manuscript, and to Laura Giovanelli and Yan Xue for technical assistance. We thank Robert I. White, Jr., for discussions of the rate of growth of pulmonary arteriovenous malformations.

Above all, we thank the family members with HHT, the Telangiectasia Self-Help Group (UK), and the HHT Foundation, International, without whom this work could not take place. This work is supported by NIH grant HL49171 to D.A.M., who is also an Established Investigator of the American Heart Association. J.N.B. is supported by a Mrs. Jean V. Baxter Medical Research Fellowship awarded by the Scottish Hospital Endowment Research Trust. T.T.S is supported by NIH fellowship 1F32HL09394, and D.W.J. is supported by NIH fellowship 1F32HL09349.

References

Attisano L, Carcamo J, Ventura F, Weis FMB, Massague J, Wrana JL (1993) Identification of human activin and TGF-β type I receptors that form heteromeric kinase complexes with type II receptors. Cell 75:671–680

Berg JN, Guttmacher AE, Marchuk DA, Porteous MEM (1996) Clinical heterogeneity in hereditary hemorrhagic telangiectasia: are pulmonary arteriovenous malformations more common in families linked to endoglin? J Med Genet 33:256–257

Braverman IM, Keh A, Jacobson BS (1990) Ultrastructure and three-dimensional organization of the telangiectases of hereditary hemorrhagic telangiectasia. J Invest Dermatol 95: 422–427

Cheifetz S, Bellon T, Cales C, Vera S, Bernabeu C, Massague J, Letarte M (1992) Endoglin is a component of the transforming growth factor-β receptor system in human endothelial cells. J Biol Chem 267:19027–19030

Ebner R, Chen RH, Shum L, Lawler S, Zioncheck TF, Lee A, Lopez AR, et al (1993) Cloning of a type I TGF-β receptor and its effect on TGF-β binding to the type II receptor. Science 260:344–348

Feng XH, Filvaroff EH, Derynck R (1995) Transforming growth factor-β (TGF-β)-induced down-regulation of cyclin A expression requires a functional TGF-β receptor complex. J Biol Chem 270:24237–24245

Guttmacher AE, Marchuk DA, White RI (1995) Current concepts: hereditary hemorrhagic telangiectasia. N Engl J Med 333:918–924

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, et al (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339

Heutink P, Haitjema T, Breedveld GJ, Janssen B, Sandkuijl LA, Bontekoe CJM, Westerman CJJ, et al (1994) Linkage of hereditary hemorrhagic telangiectasia to chromosome 9q34 and evidence for locus heterogeneity. J Med Genet 31:933–936

Ioannou PA, Amemiya CT, Garnes J, Kroisel PM, Hiroaki S, Chen C, Batzer MA, et al (1994) A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. Nat Genet 6:84–89

Johnson DW, Berg JN, Baldwin MA, Gallione CJ, Marondel I, Yoon SJ, Stenzel TT, et al (1996) Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. Nat Genet 13:189–195

Johnson DW, Berg JN, Gallione CJ, McAllister KA, Warner JP, Helmbold EA, Markel DS, et al (1995) A second locus for Hereditary Hemorrhagic Telangiectasia maps to chromosome 12. Genome Res 5:21–28

Madri JA, Kocher O, Merwin JR, Bell L, Tucker A, Basson CT (1989) Interactions of vascular cells with transforming growth factor beta. Ann NY Acad Sci 593:243–258

Masakazu T, Akioka K, Yasuda M, Ikuno Y, Oku H, Takeuchi K, Takeda T (1988) Case report: hereditary hemorrhagic telangiectasia with growing pulmonary arteriovenous fistulas followed for 24 years. Am J Med Sci 295:545–547

McAllister KA, Baldwin MA, Thukkani AK, Gallione CJ, Berg JN, Porteous ME, Guttmacher AE, et al (1995) Six novel mutations in the endoglin gene in hereditary hemorrhagic telangiectasia type 1 suggest a dominant-negative effect of receptor function. Hum Mol Genet 4:1983–1985

McAllister KA, Grogg KM, Johnson DW, Gallione CJ, Baldwin MA, Jackson CE, Helmbold EA, et al (1994*a*) Endoglin, a TGF-β receptor of endothelial cells, is the gene for hereditary hemorrhagic telangiectasia type 1 (OWR1) Nat Genet 8:345–351

McAllister KA, Lennon F, Bowles-Biesecker B, McKinnon WC, Helmbold EA, Markel DS, Jackson CE, et al (1994b) Genetic heterogeneity in hereditary haemorrhagic telangiectasia: possible correlation with clinical phenotype. J Med Genet 31:927–932

McDonald MT, Papenberg KA, Ghosh S, Glatfelter AA, Biesecker BB, Helmbold EA, Markel DS, et al (1994) A disease locus for hereditary haemorrhagic telangiectasia maps to chromosome 9q33-34. Nat Genet 6:197–204

Merwin JR, Anderson J, Kocher O, van Itallie C, Madri JA (1990) Transforming growth factor β1 modulates extracellular matrix organization and cell-cell junctional complex formation during in vitro angiogenesis. J Cell Physiol 142: 117–128

Niehrs C (1996) Growth factors: mad connection to the nucleus. Nature 381:561–562

Piantanida M, Buscarini E, Dellavecchia C, Minelli A, Rossi A, Buscarini L, Danesino C (1996) Hereditary haemorrhagic telangiectasia with extensive liver involvement is not caused by either HHT1 or HHT2. J Med Genet 33:441–443

Plauchu H, de Chadarevian J-P, Bideau A, Robert J-M (1989) Age related clinical profile of hereditary hemorrhagic telangiectasia in an epidemiologically recruited population. Am J Med Genet 32:291–297

Porteous MEM, Curtis A, Williams O, Marchuk D, Bhattacharya SS, Burn J (1994) Genetic heterogeneity in hereditary haemorrhagic telangeiectasia. J Med Genet 31:925–926

Qian F, Watnick TJ, Onuchic LF, Germino GG (1996) The molecular basis of focal cyst formation in human autosomal dominant polycystic kidney disease type 1. Cell 87:979–987

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

Schmitt J, Mielke R, Schrewe H (1995) Genomic organization of a mouse type 1 activin receptor. Biochem Biophys Res Commun 213:211–217

Senapathy P, Shapiro MB, Harris NL (1990) Splice junctions, branch point sites, and exons: sequence statistics, identifica-

tion, and applications to genome project. Methods Enzymol 183:252–278

Shovlin C, Hughes JMB, Tuddenham EGD, Temperley I, Perembelon YFN, Scott J, Seidman CE, et al (1994) A gene for hereditary hemorrhagic telangiectasia maps to chromosome 9q3. Nat Genet 6:205–209

ten Dijke P, Ichijo H, Franzen P, Schulz P, Saras J, Toyoshima H, Heldin CH, et al (1993) Activin receptor-like kinases: a novel subclass of cell-surface receptors with predicted serine/threonine kinase activity. Oncogene 8:2879–2887

Vase P, Holm M, Arendrup H (1985) Pulmonary arteriovenous fistulas in hereditary hemorrhagic telangiectasia. Acta Med Scand 218:105–109

Vincent P, Plauchu H, Hazan J, Faure S, Weissenbach J, Godet J (1995) A third locus for hereditary haemorrhagic telangiectasia maps to chromosome 12q. Hum Mol Genet 4:945–949

Yamamoto H, Ueno H, Ooshima A, Takeshita A (1996) Adenovirus-mediated transfer of a truncated transforming growth factor-β (TGF-β) type II receptor completely and specifically abolishes diverse signaling by TGF-β in vascular wall cells in primary culture. J Biol Chem 271:16253–16259

Zhang Y, Feng XH, Wu RY, Derynck, R (1996) Receptor-associated Mad homologues synergize as effectors of the TGF-β response. Nature 383:168–172